

(51) Int. Cl. 7	識別記号	F I	マーク	(参考)
G06F 15/00	310	G06F 15/00	310	A
12/00	546	12/00	546	K
13/00	520	13/00	520	C
H04N 7/173	610	H04N 7/173	610	Z

審査請求 有 請求項の数39 O L (全32頁)

(21)出願番号	特願2000-304433(P 2000-304433)
(22)出願日	平成12年10月4日(2000.10.4)
(31)優先権主張番号	特願平11-318318
(32)優先日	平成11年11月9日(1999.11.9)
(33)優先権主張国	日本 (JP)

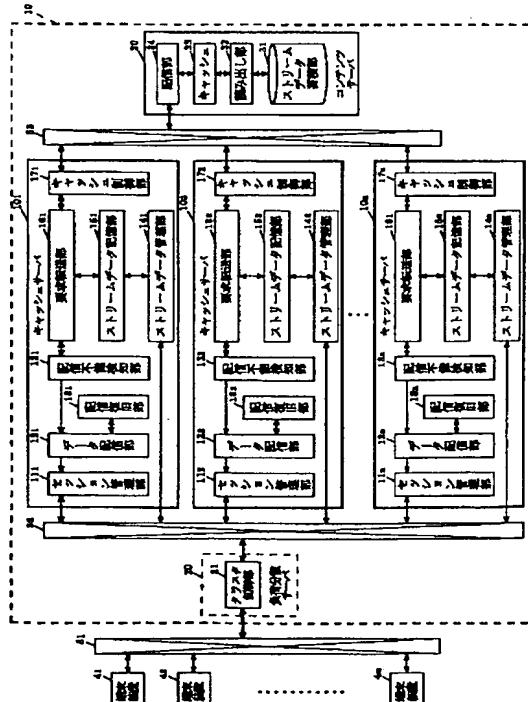
(71)出願人	000005821 松下電器産業株式会社 大阪府門真市大字門真1006番地
(72)発明者	堀内 優希 大阪府門真市大字門真1006番地 松下電器 産業株式会社内
(72)発明者	浅井 理恵子 大阪府門真市大字門真1006番地 松下電器 産業株式会社内
(72)発明者	大村 猛 大阪府門真市大字門真1006番地 松下電器 産業株式会社内
(74)代理人	100098291 弁理士 小笠原 史朗

## (54)【発明の名称】クラスタサーバ装置

## (57)【要約】

【課題】複数のキャッシュサーバへの負荷分散を最適に行いながら、いずれかのキャッシュサーバに障害が発生した場合でも、端末装置へのデータ配信が継続可能なクラスタサーバ装置を提供する。

【解決手段】クラスタ制御部は、キャッシュサーバの負荷に応じて端末装置からの要求を分散させる。キャッシュサーバは、要求されたデータ(ストリームデータ)がストリームデータ記憶部に存在する場合は当該データを、存在しない場合はコンテンツサーバから配信されるデータを、端末装置に配信する。ここで、コンテンツサーバから配信されるデータは、2つ以上のキャッシュサーバのストリームデータ記憶部に重複して記憶される。また、各々のキャッシュサーバは、データが重複して記憶されている他のキャッシュサーバの配信状態を検知しており、当該他のキャッシュサーバが配信不能になった場合には、代わってデータの配信を継続する。



## 【特許請求の範囲】

【請求項1】 ネットワークを介して接続された複数の端末装置からの要求に応じて、動画データ等の時間的に連続するストリームデータの配信を行うクラスタサーバ装置であって、

前記端末装置へ、対応するストリームデータをそれぞれ配信する複数のキャッシュサーバと、

前記端末装置から配信要求を受信し、当該配信要求に対応するストリームデータの複数の前記キャッシュサーバにおける記憶状況（記憶されているか否か）、および当該複数のキャッシュサーバがそれぞれ配信しているストリームデータの配信状態に応じて、当該配信要求をいずれかのキャッシュサーバに振り分けて転送する負荷分散サーバと、

ストリームデータを蓄積するストリームデータ蓄積部を有し、前記キャッシュサーバからの配信要求に応じて、当該配信要求に対応するストリームデータを、当該ストリームデータ蓄積部から対応する前記キャッシュサーバへ配信する1台以上のコンテンツサーバとを備え、

前記キャッシュサーバは、

前記コンテンツサーバから配信されたストリームデータを記憶するストリームデータ記憶部と、

前記負荷分散サーバから転送される配信要求を受信し、当該配信要求に対応するストリームデータが、前記ストリームデータ記憶部に記憶されている場合は、当該記憶されているストリームデータを配信するように、前記ストリームデータ記憶部に記憶されていない場合は、配信要求を前記コンテンツサーバへ転送し、前記コンテンツサーバから配信されたストリームデータを前記ストリームデータ記憶部に記憶すると共に配信する指示を出力する要求転送部と、

前記要求転送部から出力される指示に従って、前記ストリームデータ記憶部に記憶された前記ストリームデータを、前記配信要求を行った端末装置へ配信するデータ配信部と、

前記コンテンツサーバから配信されたストリームデータが、予め定めた他の前記キャッシュサーバの前記ストリームデータ記憶部にも重複して記憶されるように制御するキャッシュ制御部とを備える、クラスタサーバ装置。

## 【請求項2】 前記キャッシュサーバは、

自己が配信しているストリームデータの配信状態を管理するセッション管理部と、

前記ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、

前記負荷分散サーバは、前記端末装置から配信要求を受信した時に、前記キャッシュサーバの配信状態と当該配信要求に対応するストリームデータの記憶状況とを、全ての前記セッション管理部および前記ストリームデータ管理部に問い合わせ、その結果に基づいて、いずれかの

キャッシュサーバに当該配信要求を転送するクラスタ制御部を備え、

前記クラスタ制御部は、前記対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、前記対応するストリームデータを記憶しているキャッシュサーバがない場合は、全ての前記キャッシュサーバの中で最も負荷が低いキャッシュサーバに、前記配信要求を転送することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項3】 前記キャッシュサーバは、自己が配信しているストリームデータの配信状態を管理するセッション管理部と、

前記ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、

前記負荷分散サーバは、全ての前記キャッシュサーバの前記セッション管理部と常時通信し、各キャッシュサーバの負荷状態を保持する負荷状態管理部と、

前記端末装置から配信要求を受信した時に、当該配信要求に対応するストリームデータの記憶状況を、全ての前記ストリームデータ管理部に問い合わせ、その結果に基づいて、いずれかのキャッシュサーバに当該配信要求を転送するクラスタ制御部とを備え、

前記クラスタ制御部は、前記負荷状態管理部の情報に基づいて、前記対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、前記対応するストリームデータを記憶しているキャッシュサーバがない場合は、全ての前記キャッシュサーバの中で最も負荷が低いキャッシュサーバに、前記配信要求を転送することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項4】 前記キャッシュサーバは、自己が配信しているストリームデータの配信状態を管理するセッション管理部と、

前記ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、

前記負荷分散サーバは、全ての前記セッション管理部および前記ストリームデータ管理部と常時通信しており、その通信内容に基づいて、いずれかのキャッシュサーバに当該配信要求を転送するクラスタ制御部を備え、

前記クラスタ制御部は、前記端末装置から配信要求を受信した時に、前記対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、前記対応するストリームデータを記憶しているキャッシュサーバがない場合は、全ての前記キャッシュサ

ーバの中で最も負荷が低いキャッシングサーバに、当該配信要求を転送することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項5】 第1～第nのキャッシングサーバ(nは、2以上の整数)で構成される場合、

第iのキャッシングサーバ(iは、1～nのいずれか)において、前記キャッシング制御部は、前記ストリームデータ記憶部にストリームデータを記憶させる場合に、第

(i+1)のキャッシングサーバ(ただし、i=nの場合は第1のキャッシングサーバ)の前記ストリームデータ記憶部に同じストリームデータを記憶するように制御することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項6】 前記キャッシング制御部は、配信要求に対応するストリームデータが前記ストリームデータ記憶部にない場合、前記コンテンツサーバへ転送する配信要求の送信先アドレスを、予め設定されたマルチキャストアドレスに変更して送信し、前記予め設定されたマルチキャストアドレス宛のストリームデータを受信した場合には、前記ストリームデータ記憶部に記憶するように制御することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項7】 前記第iのキャッシングサーバにおいて、前記キャッシング制御部は、配信要求に対応するストリームデータが前記ストリームデータ記憶部にない場合、当該配信要求と共に、送信先アドレスに前記第(i+1)のキャッシングサーバのアドレスを指定した配信要求を、前記コンテンツサーバへ送信することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項8】 前記クラスタ制御部は、1台のキャッシングサーバの配信ストリーム数が、値{( (n-1)/n) × MAX}以下であって(MAXは、前記キャッシングサーバ1台当たりの配信可能な最大配信ストリーム性能)、前記第iのキャッシングサーバおよび第(i+1)のキャッシングサーバの前記ストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計が、MAX以下に制限されるように、前記配信要求を前記キャッシングサーバに振り分けて転送することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項9】 前記クラスタ制御部は、前記第iのキャッシングサーバおよび第(i+1)のキャッシングサーバの前記ストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計、および、前記第(i-1)のキャッシングサーバ(ただし、i=1の場合は第nのキャッシングサーバ)および第iのキャッシングサーバの前記ストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計が、それぞれMAX(MAXは、前記キャッシングサーバ1台当たりの配信可能な最大配信ストリーム性能)以

下に制限されるように、前記配信要求を前記キャッシングサーバに振り分けて転送することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項10】 前記キャッシングサーバは、前記データ配信部から、重複して同じストリームデータを記憶する前記他のキャッシングサーバに、前記端末装置へ現在配信している配信情報を、一定時間間隔で通知しており、

10 前記他のキャッシングサーバから通知される各々の前記配信情報によって、いずれか1台以上の当該他のキャッシングサーバが配信不能になったことを検知した場合、ストリームデータが途切れないよう定められた時間以内に、配信不能になった当該他のキャッシングサーバに代わって、当該他のキャッシングサーバが最後に配信したストリームデータの続きを配信するよう前記データ配信部へ指示する配信不能検知部をさらに備える、請求項1に記載のクラスタサーバ装置。

【請求項11】 前記第iのキャッシングサーバは、前記データ配信部から、重複して同じストリームデータを記憶する前記第(i+1)のキャッシングサーバおよび第(i-1)のキャッシングサーバに、前記端末装置へ現在配信している配信情報を、一定時間間隔で通知しており、

20 前記第(i+1)のキャッシングサーバおよび第(i-1)のキャッシングサーバから通知される各々の前記配信情報によって、いずれか一方または双方のキャッシングサーバが配信不能になったことを検知した場合、ストリームデータが途切れないよう定められた時間以内に、当該配信不能になったキャッシングサーバに代わって、当該配信不能になったキャッシングサーバが最後に配信したストリームデータの続きを配信するよう前記データ配信部へ指示する配信不能検知部をさらに備える、請求項5に記載のクラスタサーバ装置。

【請求項12】 前記配信不能検知部の指示に従って、前記データ配信部が、配信不能となったキャッシングサーバに代わって前記端末装置へストリームデータを配信している場合、

前記キャッシング制御部は、前記ストリームデータ記憶部に新たなストリームデータを記憶しないように制御することを特徴とする、請求項10に記載のクラスタサーバ装置。

【請求項13】 前記配信不能検知部の指示に従って、前記データ配信部が、配信不能となったキャッシングサーバに代わって前記端末装置へストリームデータを配信している場合、

前記キャッシング制御部は、前記ストリームデータ記憶部に新たなストリームデータを記憶しないように制御することを特徴とする、請求項11に記載のクラスタサーバ装置。

【請求項14】 前記配信不能となっていたキャッシング

サーバが復旧した場合、

前記復旧したキャッシュサーバは、重複して同じストリームデータを記憶する他の前記キャッシュサーバの前記データ配信部へ復旧通知を出力し、  
配信不能時に、前記復旧したキャッシュサーバに代わって前記端末装置へストリームデータを配信していた前記他のキャッシュサーバは、前記復旧通知を受けて、前記ストリームデータ記憶部に記憶していたストリームデータを、前記復旧したキャッシュサーバの前記ストリームデータ記憶部にコピーすることを特徴とする、請求項10に記載のクラスタサーバ装置。

【請求項15】 前記配信不能となっていたキャッシュサーバが復旧した場合、

前記復旧したキャッシュサーバは、重複して同じストリームデータを記憶する他の前記キャッシュサーバの前記データ配信部へ復旧通知を出力し、  
配信不能時に、前記復旧したキャッシュサーバに代わって前記端末装置へストリームデータを配信していた前記他のキャッシュサーバは、前記復旧通知を受けて、前記ストリームデータ記憶部に記憶していたストリームデータを、前記復旧したキャッシュサーバの前記ストリームデータ記憶部にコピーすることを特徴とする、請求項11に記載のクラスタサーバ装置。

【請求項16】 前記キャッシュサーバは、

前記端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションと、それぞれのセッションに関して前記データ配信部が最後に送信したパケット（ストリームデータの一単位）と、ストリームデータの配信ストリーム数とを、管理するセッション管理部と、

前記セッション管理部で管理されている全てのセッションのパケットを、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知する配信不能検知部とをさらに備え、

前記配信不能検知部は、前記通知が一定時間来ない場合、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないと定められた時間以内に、当該配信不能になったキャッシュサーバに代わって、同じストリームデータを記憶するいずれか1台の他のキャッシュサーバが、当該配信不能になったキャッシュサーバで管理される全てのセッションに関して、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するよう前記データ配信部へ指示することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項17】 前記キャッシュサーバは、

前記端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションと、それぞれのセッションに関して前記データ配信部が最後に送信したパケット（ストリームデータの一単位）と、ストリームデータ

10

タの配信ストリーム数とを、管理するセッション管理部と、  
前記セッション管理部で管理されている全てのセッションのパケットを、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知する配信不能検知部とをさらに備え、  
前記配信不能検知部は、前記通知が一定時間来ない場合、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないと定められた時間以内に、当該配信不能になったキャッシュサーバに代わって、同じストリームデータを記憶するいずれか1台の他のキャッシュサーバが、当該配信不能になったキャッシュサーバで管理される全てのセッションに関して、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するよう前記データ配信部へ指示することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項18】 前記キャッシュサーバは、  
前記端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションを識別するためのセッション識別子と、それぞれのセッションに関して前記データ配信部が最後に送信したパケット（ストリームデータの一単位）を識別するためのパケット識別子と、当該セッション識別子およびパケット識別子を格納するためのセッション管理テーブルと、当該セッション管理テーブルを当該セッション識別子の値によって2つに分類するためのテーブル境界値と、ストリームデータの配信ストリーム数とを、管理するセッション管理部と、

前記セッション管理部で管理されているストリームデータを配信中の全てのセッションの前記セッション識別子および前記パケット識別子を、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知し、当該他のキャッシュサーバから受け取る通知に従って、前記セッション管理テーブルに格納されている前記セッション識別子および前記パケット識別子の値を更新する、配信不能検知部とをさらに備え、  
前記セッション管理部は、自己の前記データ配信部が配信しているストリームデータのセッションに関しては、

40 前記セッション識別子を前記テーブル境界値より小さい値に設定し、前記他のキャッシュサーバの前記データ配信部が配信しているストリームデータのセッションに関しては、当該他のキャッシュサーバで管理される前記セッション識別子の値を、前記テーブル境界値より大きい値に変換して設定し、  
前記データ配信部は、全てのセッションにおいて、前記セッション識別子の値が前記テーブル境界値より小さい場合は、パケットを前記端末装置へ送信すると共に、前記パケット識別子の値を更新することを繰り返し、

前記配信不能検知部は、前記通知が一定時間来ない場合

は、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないよう定められた時間以内に、前記テーブル境界値を前記セッション識別子の最大値より大きい値に設定することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項19】 前記キャッシュサーバは、前記端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションを識別するためのセッション識別子と、それぞれのセッションに関して前記データ配信部が最後に送信したパケット（ストリームデータの一単位）を識別するためのパケット識別子と、当該セッション識別子およびパケット識別子を格納するためのセッション管理テーブルと、当該セッション管理テーブルを当該セッション識別子の値によって2つに分類するためのテーブル境界値と、ストリームデータの配信ストリーム数とを、管理するセッション管理部と、前記セッション管理部で管理されているストリームデータを配信中の全てのセッションの前記セッション識別子および前記パケット識別子を、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知し、当該他のキャッシュサーバから受け取る通知に従って、前記セッション管理テーブルに格納されている前記セッション識別子および前記パケット識別子の値を更新する、配信不能検知部とをさらに備え、

前記セッション管理部は、自己の前記データ配信部が配信しているストリームデータのセッションに関しては、前記セッション識別子を前記テーブル境界値より小さい値に設定し、前記他のキャッシュサーバの前記データ配信部が配信しているストリームデータのセッションに関しては、当該他のキャッシュサーバで管理される前記セッション識別子の値を、前記テーブル境界値より大きい値に変換して設定し、

前記データ配信部は、全てのセッションにおいて、前記セッション識別子の値が前記テーブル境界値より小さい場合は、パケットを前記端末装置へ送信すると共に、前記パケット識別子の値を更新することを繰り返し、

前記配信不能検知部は、前記通知が一定時間来ない場合は、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないよう定められた時間以内に、前記テーブル境界値を前記セッション識別子の最大値より大きい値に設定することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項20】 前記キャッシュサーバは、自己が配信しているストリームデータの配信ストリーム数を管理するセッション管理部と、他の前記キャッシュサーバが配信不能になった場合、当該配信不能になった他のキャッシュサーバを検知し、当該他のキャッシュサーバが配信していた全てのストリームデータの配信を、当該他のキャッシュサーバに代わつ

て継続することを前記データ配信部へ指示する配信不能検知部とをさらに備え、

前記配信不能検知部は、前記セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、同じストリームデータを記憶するキャッシュサーバのうち、前記配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するよう前記データ配信部へ指示し、

前記あふれ通知を受け取ったさらに他のキャッシュサーバにおける前記配信不能検知部は、前記あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリームデータの全てについて、前記あふれ通知を出力したキャッシュサーバが最後に配信したストリームデータの続きを配信するよう前記データ配信部へ指示することを特徴とする、請求項1に記載のクラスタサーバ装置。

【請求項21】 前記キャッシュサーバは、自己が配信しているストリームデータの配信ストリーム数を管理するセッション管理部と、

他の前記キャッシュサーバが配信不能になった場合、当該配信不能になった他のキャッシュサーバを検知し、当該他のキャッシュサーバが配信していた全てのストリームデータの配信を、当該他のキャッシュサーバに代わつて継続することを前記データ配信部へ指示する配信不能検知部とをさらに備え、

前記配信不能検知部は、前記セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、同じストリームデータを記憶するキャッシュサーバのうち、前記配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するよう前記データ配信部へ指示し、

前記あふれ通知を受け取ったさらに他のキャッシュサーバにおける前記配信不能検知部は、前記あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリームデータの全てについて、前記あふれ通知を出力したキャッシュサーバが最後に配信したストリームデータの続きを配信するよう前記データ配信部へ指示することを特徴とする、請求項5に記載のクラスタサーバ装置。

【請求項22】 他の前記キャッシュサーバが配信不能になった場合であって、前記セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、前記配信不能検知部は、同じストリームデータを記憶するキャッシュサーバのうち、前記配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力

すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するよう前記データ配信部へ指示し、

前記あふれ通知を受け取ったさらに他のキャッシュサーバにおける前記配信不能検知部は、前記あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリームデータ全てについて、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するよう前記データ配信部へ指示することを特徴とする、請求項16に記載のクラスタサーバ装置。

【請求項23】他の前記キャッシュサーバが配信不能になった場合であって、前記セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、前記配信不能検知部は、同じストリームデータを記憶するキャッシュサーバのうち、前記配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するよう前記データ配信部へ指示し、

前記あふれ通知を受け取ったさらに他のキャッシュサーバにおける前記配信不能検知部は、前記あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリームデータ全てについて、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するよう前記データ配信部へ指示することを特徴とする、請求項17に記載のクラスタサーバ装置。

【請求項24】同じストリームデータを記憶する前記キャッシュサーバは、最大セッション情報登録数がMAX (MAXは、前記キャッシュサーバ1台当たりの配信可能な最大配信ストリーム性能) の前記セッション管理テーブルをそれぞれ有し、

それぞれの前記セッション管理部は、自己の前記データ配信部が配信しているストリームデータのセッションについては、前記セッション識別子を前記テーブル境界値より小さい値に設定し、他の前記キャッシュサーバの前記データ配信部が配信しているストリームデータのセッションについては、当該他のキャッシュサーバの前記セッション管理テーブルにおいて、前記テーブル境界値より小さいセッション識別子の値I1～Inを値(MAX-I1-1)～(MAX-In-1)に変換して設定し、

前記配信不能検知部は、前記他のキャッシュサーバから受け取る通知に従って、当該通知に含まれるセッション識別子I1～Inを(MAX-I1-1)～(MAX-In-1)に変換した値と、パケット識別子P1～Pnの値とを、前記セッション管理テーブルの値として更新し、通知される前記テーブル境界値がFに変更された場

合、前記テーブル境界値を(MAX-F)に設定することを特徴とする、請求項18に記載のクラスタサーバ装置。

【請求項25】同じストリームデータを記憶する前記キャッシュサーバは、最大セッション情報登録数がMAX (MAXは、前記キャッシュサーバ1台当たりの配信可能な最大配信ストリーム性能) の前記セッション管理テーブルをそれぞれ有し、

それぞれの前記セッション管理部は、自己の前記データ配信部が配信しているストリームデータのセッションについては、前記セッション識別子を前記テーブル境界値より小さい値に設定し、他の前記キャッシュサーバの前記データ配信部が配信しているストリームデータのセッションについては、当該他のキャッシュサーバの前記セッション管理テーブルにおいて、前記テーブル境界値より小さいセッション識別子の値I1～Inを値(MAX-I1-1)～(MAX-In-1)に変換して設定し、

前記配信不能検知部は、前記他のキャッシュサーバから受け取る通知に従って、当該通知に含まれるセッション識別子I1～Inを(MAX-I1-1)～(MAX-In-1)に変換した値と、パケット識別子P1～Pnの値とを、前記セッション管理テーブルの値として更新し、通知される前記テーブル境界値がFに変更された場合、前記テーブル境界値を(MAX-F)に設定することを特徴とする、請求項19に記載のクラスタサーバ装置。

【請求項26】キャッシュサーバAにおいて、キャッシュサーバAとキャッシュサーバBとの両方に記憶されるストリームデータ集合S1と、キャッシュサーバAとキャッシュサーバCとの両方に記憶されるストリームデータ集合S2とを、前記ストリームデータ記憶部に記憶する場合、

前記キャッシュサーバAが、セッション管理テーブルT0, T1と、セッション管理テーブルT0, T1を前記セッション識別子の値によって2つに分類するためのテーブル境界値F0, F1とを有し、

前記セッション管理テーブルT0には、前記キャッシュサーバAが配信中の前記ストリームデータ集合S1に含まれるストリームデータのセッションについては、前記セッション識別子が前記テーブル境界値F0より小さい値に設定され、前記キャッシュサーバBが配信中の前記ストリームデータ集合S1に含まれるストリームデータのセッションについては、前記キャッシュサーバBの前記セッション管理テーブルで、前記セッション識別子が前記テーブル境界値F0より小さい値I1～Inを値(MAX-I1-1)～(MAX-In-1)に変換されて設定され、

前記セッション管理テーブルT1には、前記キャッシュサーバAが配信中の前記ストリームデータ集合S2に含

まれるストリームデータのセッションに関しては、前記セッション識別子が前記テーブル境界値F1より小さい値に設定され、前記キャッシュサーバCが配信中の前記ストリームデータ集合S2に含まれるストリームデータのセッションに関しては、前記キャッシュサーバCの前記セッション管理テーブルで、前記セッション識別子が前記テーブル境界値F1より小さい値J1～Jnを値 $(MAX - J1 - 1) \sim (MAX - Jn - 1)$ に変換されて設定され、

前記キャッシュサーバAにおいて、前記配信不能検知部は、

前記セッション管理部に問い合わせて、前記キャッシュサーバB, Cと互いにストリームデータを配信中の全てのセッションについて、前記セッション識別子および前記パケット識別子を通知し、

前記キャッシュサーバBからの通知が一定時間来ない場合は、前記テーブル境界値F0を、前記キャッシュサーバCからの通知が一定時間来ない場合は、前記テーブル境界値F1を、MAX以上に設定し、

ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合であって、前記キャッシュサーバBからの通知が来ない場合は、前記テーブル境界値F1を、前記キャッシュサーバCからの通知が来ない場合は、前記テーブル境界値F0を、キャッシュサーバのストリーム配信数がキャッシュサーバの最大配信ストリーム性能以下になる値に変更することを特徴とする、請求項18に記載のクラスタサーバ装置。

【請求項27】 キャッシュサーバAにおいて、キャッシュサーバAとキャッシュサーバBとの両方に記憶されるストリームデータ集合S1と、キャッシュサーバAとキャッシュサーバCとの両方に記憶されるストリームデータ集合S2とを、前記ストリームデータ記憶部に記憶する場合、

前記キャッシュサーバAが、セッション管理テーブルT0, T1と、セッション管理テーブルT0, T1を前記セッション識別子の値によって2つに分類するためのテーブル境界値F0, F1とを有し、

前記セッション管理テーブルT0には、前記キャッシュサーバAが配信中の前記ストリームデータ集合S1に含まれるストリームデータのセッションに関しては、前記セッション識別子が前記テーブル境界値F0より小さい値に設定され、前記キャッシュサーバBが配信中の前記ストリームデータ集合S1に含まれるストリームデータのセッションに関しては、前記キャッシュサーバBの前記セッション管理テーブルで、前記セッション識別子が前記テーブル境界値F0より小さい値I1～Inを値 $(MAX - I1 - 1) \sim (MAX - In - 1)$ に変換されて設定され、

前記セッション管理テーブルT1には、前記キャッシュサーバAが配信中の前記ストリームデータ集合S2に含

まれるストリームデータのセッションに関しては、前記セッション識別子が前記テーブル境界値F1より小さい値に設定され、前記キャッシュサーバCが配信中の前記ストリームデータ集合S2に含まれるストリームデータのセッションに関しては、前記キャッシュサーバCの前記セッション管理テーブルで、前記セッション識別子が前記テーブル境界値F1より小さい値J1～Jnを値 $(MAX - J1 - 1) \sim (MAX - Jn - 1)$ に変換されて設定され、

10 前記キャッシュサーバAにおいて、前記配信不能検知部は、

前記セッション管理部に問い合わせて、前記キャッシュサーバB, Cと互いにストリームデータを配信中の全てのセッションについて、前記セッション識別子および前記パケット識別子を通知し、

前記キャッシュサーバBからの通知が一定時間来ない場合は、前記テーブル境界値F0を、前記キャッシュサーバCからの通知が一定時間来ない場合は、前記テーブル境界値F1を、MAX以上に設定し、

20 ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合であって、前記キャッシュサーバBからの通知が来ない場合は、前記テーブル境界値F1を、前記キャッシュサーバCからの通知が来ない場合は、前記テーブル境界値F0を、キャッシュサーバのストリーム配信数がキャッシュサーバの最大配信ストリーム性能以下になる値に変更することを特徴とする、請求項19に記載のクラスタサーバ装置。

【請求項28】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータの続きを配信するよう前記データ配信部へ指示し、

40 前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項14に記載のクラスタサーバ装置。

【請求項29】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータの続きを配信するよう前記データ配信部へ指示し、

ムデータを配信していた前記キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータの続きを配信するよう前記データ配信部へ指示し、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項 15 に記載のクラスタサーバ装置。

【請求項 30】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、前記キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット識別子で識別されるパケットの次のパケットから配信を継続するよう前記データ配信部へ指示し、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項 16 に記載のクラスタサーバ装置。

【請求項 31】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、前記キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット

識別子で識別されるパケットの次のパケットから配信を継続するよう前記データ配信部へ指示し、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項 17 に記載のクラスタサーバ装置。

【請求項 32】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、前記キャッシュサーバへ復旧通知を出力すると共に、前記テーブル境界値を配信不能になる前の値に設定し、

前記キャッシュサーバの前記配信復旧部は、前記テーブル境界値を、前記復旧した他のキャッシュサーバが配信不能になる前の値に設定することを特徴とする、請求項 24 に記載のクラスタサーバ装置。

【請求項 33】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部がストリームデータの復旧を完了した後、前記キャッシュサーバへ復旧通知を出力すると共に、前記テーブル境界値を配信不能になる前の値に設定し、

前記キャッシュサーバの前記配信復旧部は、前記テーブル境界値を、前記復旧した他のキャッシュサーバが配信不能になる前の値に設定することを特徴とする、請求項 25 に記載のクラスタサーバ装置。

【請求項 34】 前記配信不能になった他のキャッシュ

サーバの代わりにストリームデータを配信していたことで、前記キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示した後、前記復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータの続きを配信するよう前記データ配信部へ指示し、

前記さらに他のキャッシュサーバの前記配信復旧部は、前記復旧通知を出力してきた前記キャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項28に記載のクラスタサーバ装置。

【請求項35】 前記配信不能になった他のキャッシュサーバの代わりにストリームデータを配信していたことで、前記キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示した後、前記復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータの続きを配信するよう前記データ配信部へ指示し、

前記さらに他のキャッシュサーバの前記配信復旧部は、前記復旧通知を出力してきた前記キャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項29に記載のクラスタサーバ装置。

【請求項36】 前記配信不能になった他のキャッシュサーバの代わりにストリームデータを配信していたことで、前記キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、

前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示した後、前記復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット識別子で識別されるパケットの次のパケットから配信を継続するよう前記データ配信部へ指示し、

前記さらに他のキャッシュサーバの前記配信復旧部は、

前記復旧通知を出力してきた前記キャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項30に記載のクラスタサーバ装置。

【請求項37】 前記配信不能になった他のキャッシュサーバの代わりにストリームデータを配信していたことで、前記キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、

10 前記キャッシュサーバの前記配信復旧部は、前記復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示した後、前記復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット識別子で識別されるパケットの次のパケットから配信を継続するよう前記データ配信部へ指示し、

前記さらに他のキャッシュサーバの前記配信復旧部は、前記復旧通知を出力してきた前記キャッシュサーバの代わりに配信していたストリームデータの配信を、停止するよう前記データ配信部へ指示することを特徴とする、請求項31に記載のクラスタサーバ装置。

【請求項38】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

前記配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバAの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部が復旧通知を行ってストリームデータの復旧を完了した後、前記テーブル境界値を配信不能になる前の値に変更し、

40 前記キャッシュサーバAの前記配信復旧部は、前記復旧したキャッシュサーバから通知される前記テーブル境界値が、前記復旧したキャッシュサーバが配信不能になる前の値に戻った場合、前記テーブル境界値F0, F1を前記復旧したキャッシュサーバが配信不能になる前の値に戻すことを特徴とする、請求項26に記載のクラスタサーバ装置。

【請求項39】 前記キャッシュサーバは、前記配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、

50 前記配信不能となっていた他のキャッシュサーバが復旧

した場合、当該復旧した他のキャッシュサーバの前記キャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバの前記データ配信部へ復旧通知を出力し、代わりに前記端末装置へストリームデータを配信していた前記キャッシュサーバAの前記ストリームデータ記憶部に記憶されているストリームデータを、前記復旧した他のキャッシュサーバの前記ストリームデータ記憶部にコピーし、

前記復旧した他のキャッシュサーバの前記配信復旧部は、前記キャッシュ制御部が復旧通知を行ってストリームデータの復旧を完了した後、前記テーブル境界値を配信不能になる前の値に変更し、

前記キャッシュサーバAの前記配信復旧部は、前記復旧したキャッシュサーバから通知される前記テーブル境界値が、前記復旧したキャッシュサーバが配信不能になる前の値に戻った場合、前記テーブル境界値F0, F1を前記復旧したキャッシュサーバが配信不能になる前の値に戻すことを特徴とする、請求項27に記載のクラスタサーバ装置。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】本発明は、クラスタサーバ装置に関し、より特定的には、ネットワークで接続された複数の端末装置から受け付けるリクエストに従って、対応する映像等のストリームデータを配信するクラスタサーバ装置に関する。

##### 【0002】

【従来の技術】映像等に代表されるストリームデータの配信を行うクラスタサーバ装置には、ネットワークで接続された端末装置に、ストリームデータを途切れなく配信することが要求される。ここで、ストリームデータのデータ量は、一般的に膨大なものであるため、クラスタサーバ装置には、データ格納用に大容量のハードディスク等の記憶装置が使用されている。しかし、データ格納用に大容量の記憶装置を備えたサーバで構成されるクラスタサーバ装置では、サーバの能力の限界値によって配信可能な端末装置の数に制限がある。そこで、従来においては、配信可能な端末装置の数をより多くするため、クラスタサーバ装置内に複数のサーバを設け、端末装置から受け付けるリクエストを複数のサーバに順番に振り分ける方式等が考案されている。以下、リクエストを複数のサーバに順番に振り分ける方式による従来のクラスタサーバ装置を、簡単に説明する。

【0003】図18は、リクエストを複数のサーバに順番に振り分ける方式による従来のクラスタサーバ装置を用いた配信システムの構成の一例を示すブロック図である。図18において、従来の配信システムは、m個の端末装置401～40m (mは、2以上の整数) とクラスタサーバ装置200とが、ネットワーク501で接続されている。クラスタサーバ装置200は、n個のキャッシュサーバ201～20n (nは、2以上の整数) とコンテンツサーバ300とで構成され、キャッシュサーバ201～20nは、内部ネットワーク503によってコンテンツサーバ300にそれぞれ接続されている。キャッシュサーバ201～20nは、受信部211～21nと、キャッシュ221～22nと、配信部231～23nとを、それぞれ備えている。また、コンテンツサーバ300は、ストリームデータ蓄積部301と、読み出し部302と、キャッシュ303と、配信部304とを備えている。なお、ストリームデータは、ストリームデータ蓄積部301にブロック単位で格納されている。

【0004】今、端末装置401が、コンテンツサーバ300に対して、あるストリームデータを要求したとする。この要求に対して、コンテンツサーバ300の読み出し部302は、該当するストリームデータの読み出し要求をストリームデータ蓄積部301へ発行し、読み出したストリームデータをキャッシュ303に一旦蓄える。配信部304は、キャッシュ303に蓄えられたストリームデータをキャッシュサーバ201へ送信する。

【0005】このように、従来のクラスタサーバ装置200では、キャッシュサーバ201～20nを並列に構成して要求処理を巡回させることで、システム全体における配信可能な端末装置401～40mの数の増加を実現している。

【0006】また、複数のサーバにリクエストを順番に振り分ける方式に加え、1台のサーバに障害が発生した場合には、障害復旧後にデータのリカバリを行う従来のクラスタサーバ装置が、特開平8-263351号公報に開示されている。この公報で開示されている従来のクラスタサーバ装置では、各キャッシュサーバにコンテンツサーバの更新履歴をそれぞれ保持させておき、障害復旧後には、キャッシュサーバで保持している最新の更新履歴情報を用いて、コンテンツサーバのデータのリカバリを行う。これにより、従来のクラスタサーバ装置は、1台のコンテンツサーバのデータに障害が発生した場合でも、システムを中断させることなく、かつ迅速にリカ

リバリを行っている。

【0007】一方で、従来のクラスタサーバ装置では、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。また、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0008】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0009】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0010】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0011】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0012】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0013】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0014】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0015】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0016】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0017】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0018】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0019】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0020】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0021】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0022】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0023】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0024】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0025】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0026】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0027】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0028】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0029】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0030】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0031】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0032】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0033】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0034】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0035】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0036】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0037】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0038】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0039】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0040】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0041】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0042】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0043】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0044】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0045】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0046】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0047】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0048】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0049】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

【0050】そこで、本発明は、各キャッシュサーバの更新履歴を保持するため、データのリカバリに時間がかかる問題がある。

バリ処理を行うことを実現している。

【0007】

【発明が解決しようとする課題】しかしながら、上記従来のクラスタサーバ装置200においては、端末装置401～404から受け付けるリクエスト単位で、複数のキャッシュサーバ201～204にリクエストを巡回的に振り分けている。このため、従来のクラスタサーバ装置200では、動画データのような時間的に連続なストリームデータが配信されている場合、複数のキャッシュサーバ201～204に負荷がうまく分散されず、動画が途切れることがあった。また、ストリームデータの配信途中にキャッシュサーバに障害が発生した場合には、他のキャッシュサーバによって、当該ストリームデータを最初から配信し直さなくてはならないといった問題がある。さらに、上記公報に開示されている従来のクラスタサーバ装置は、コンテンツサーバの更新履歴をキャッシュサーバで保持しているだけなので、キャッシュサーバ自体が故障した場合には、端末装置に対するサービスが途切れてしまうという問題がある。

【0008】それ故、本発明の目的は、端末装置から受け付ける要求に対して、複数のキャッシュサーバへの負荷分散を最適に行なながら、かつ、いずれかのキャッシュサーバに障害が発生した場合でも、端末装置へのデータ配信を問題なく（動画が途切ることなく）継続でき、かつ障害復旧できるクラスタサーバ装置を提供することである。

【0009】

【課題を解決するための手段および発明の効果】第1の発明は、ネットワークを介して接続された複数の端末装置からの要求に応じて、動画データ等の時間的に連続するストリームデータの配信を行うクラスタサーバ装置であって、端末装置へ、対応するストリームデータをそれぞれ配信する複数のキャッシュサーバと、端末装置から配信要求を受信し、当該配信要求に対応するストリームデータの複数のキャッシュサーバにおける記憶状況（記憶されているか否か）、および当該複数のキャッシュサーバがそれぞれ配信しているストリームデータの配信状態に応じて、当該配信要求をいずれかのキャッシュサーバに振り分けて転送する負荷分散サーバと、ストリームデータを蓄積するストリームデータ蓄積部を有し、キャッシュサーバからの配信要求に応じて、当該配信要求に対応するストリームデータを、当該ストリームデータ蓄積部から対応するキャッシュサーバへ配信する1台以上のコンテンツサーバとを備え、キャッシュサーバは、コンテンツサーバから配信されたストリームデータを記憶するストリームデータ記憶部と、負荷分散サーバから転送される配信要求を受信し、当該配信要求に対応するストリームデータが、ストリームデータ記憶部に記憶されている場合は、当該記憶されているストリームデータを配信するように、ストリームデータ記憶部に記憶されて

いない場合は、配信要求をコンテンツサーバへ転送し、コンテンツサーバから配信されたストリームデータをストリームデータ記憶部に記憶すると共に配信する指示を出力する要求転送部と、要求転送部から出力される指示に従って、ストリームデータ記憶部に記憶されたストリームデータを、配信要求を行った端末装置へ配信するデータ配信部と、コンテンツサーバから配信されたストリームデータが、予め定めた他のキャッシュサーバのストリームデータ記憶部にも重複して記憶されるように制御するキャッシュ制御部とを備える。

【0010】第2の発明は、第1の発明に従属する発明であって、キャッシュサーバは、自己が配信しているストリームデータの配信状態を管理するセッション管理部と、ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、負荷分散サーバは、端末装置から配信要求を受信した時に、キャッシュサーバの配信状態と当該配信要求に対応するストリームデータの記憶状況とを、全てのセッション管理部およびストリームデータ管理部に問い合わせ、その結果に基づいて、いずれかのキャッシュサーバに当該配信要求を転送するクラスタ制御部を備え、クラスタ制御部は、対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、対応するストリームデータを記憶しているキャッシュサーバがない場合は、全てのキャッシュサーバの中で最も負荷が低いキャッシュサーバに、配信要求を転送することを特徴とする。

【0011】第3の発明は、第1の発明に従属する発明であって、キャッシュサーバは、自己が配信しているストリームデータの配信状態を管理するセッション管理部と、ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、負荷分散サーバは、全てのキャッシュサーバのセッション管理部と常時通信し、各キャッシュサーバの負荷状態を保持する負荷状態管理部と、端末装置から配信要求を受信した時に、当該配信要求に対応するストリームデータの記憶状況を、全てのストリームデータ管理部に問い合わせ、その結果に基づいて、いずれかのキャッシュサーバに当該配信要求を転送するクラスタ制御部とを備え、クラスタ制御部は、負荷状態管理部の情報に基づいて、対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、対応するストリームデータを記憶しているキャッシュサーバがない場合は、全てのキャッシュサーバの中で最も負荷が低いキャッシュサーバに、配信要求を転送することを特徴とする。

【0012】第4の発明は、第1の発明に従属する発明であって、キャッシュサーバは、自己が配信しているス

40 40

50

ストリームデータの配信状態を管理するセッション管理部と、ストリームデータ記憶部に記憶されているストリームデータを管理するストリームデータ管理部とをさらに備え、負荷分散サーバは、全てのセッション管理部およびストリームデータ管理部と常時通信しており、その通信内容に基づいて、いずれかのキャッシュサーバに当該配信要求を転送するクラスタ制御部を備え、クラスタ制御部は、端末装置から配信要求を受信した時に、対応するストリームデータを記憶しているキャッシュサーバがある場合は、当該記憶しているキャッシュサーバの中で最も負荷が低いキャッシュサーバに、対応するストリームデータを記憶しているキャッシュサーバがない場合は、全てのキャッシュサーバの中で最も負荷が低いキャッシュサーバに、当該配信要求を転送することを特徴とする。

【0013】第5の発明は、第1の発明に従属する発明であって、第1～第nのキャッシュサーバ（nは、2以上の整数）で構成される場合、第iのキャッシュサーバ（iは、1～nのいずれか）において、キャッシュ制御部は、ストリームデータ記憶部にストリームデータを記憶させる場合に、第（i+1）のキャッシュサーバ（ただし、i=nの場合は第1のキャッシュサーバ）のストリームデータ記憶部に同じストリームデータを記憶するように制御することを特徴とする。

【0014】第6の発明は、第5の発明に従属する発明であって、キャッシュ制御部は、配信要求に対応するストリームデータがストリームデータ記憶部にない場合、コンテンツサーバへ転送する配信要求の送信先アドレスを、予め設定されたマルチキャストアドレスに変更して送信し、予め設定されたマルチキャストアドレス宛のストリームデータを受信した場合には、ストリームデータ記憶部に記憶するように制御することを特徴とする。

【0015】第7の発明は、第5の発明に従属する発明であって、第iのキャッシュサーバにおいて、キャッシュ制御部は、配信要求に対応するストリームデータがストリームデータ記憶部にない場合、当該配信要求と共に、送信先アドレスに第（i+1）のキャッシュサーバのアドレスを指定した配信要求を、コンテンツサーバへ送信することを特徴とする。

【0016】第8の発明は、第5の発明に従属する発明であって、クラスタ制御部は、1台のキャッシュサーバの配信ストリーム数が、値 { ((n-1)/n) × MAX } 以下であって（MAXは、キャッシュサーバ1台当たりの配信可能な最大配信ストリーム性能）、第iのキャッシュサーバおよび第（i+1）のキャッシュサーバのストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計が、MAX以下に制限されるように、配信要求をキャッシュサーバに振り分けて転送することを特徴とする。

【0017】第9の発明は、第5の発明に従属する発明

であって、クラスタ制御部は、第iのキャッシュサーバおよび第（i+1）のキャッシュサーバのストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計、および、第（i-1）のキャッシュサーバ（ただし、i=1の場合は第nのキャッシュサーバ）および第iのキャッシュサーバのストリームデータ記憶部にそれぞれ記憶されているストリームデータの配信ストリーム数の合計が、それぞれMAX以下に制限されるように、配信要求をキャッシュサーバに振り分けて転送することを特徴とする。

【0018】上記のように、第1～第9の発明によれば、1つのストリームデータを複数のキャッシュサーバに記憶しておく。これにより、複数のキャッシュサーバからストリームデータを直接配信することが可能となるので、配信性能の飛躍的向上を図ることができる。また、記憶されているストリームデータに応じてキャッシュサーバの負荷を均等に分散させることで、記憶部（キャッシュ）の効率的な運用、かつ、より多くの端末装置へのデータ配信が可能になる。

【0019】第10の発明は、第1の発明に従属する発明であって、キャッシュサーバは、データ配信部から、重複して同じストリームデータを記憶する他のキャッシュサーバに、端末装置へ現在配信している配信情報を、一定時間間隔で通知しており、他のキャッシュサーバから通知される各々の配信情報によって、いずれか1台以上の当該他のキャッシュサーバが配信不能になったことを検知した場合、ストリームデータが途切れないよう定められた時間以内に、配信不能になった当該他のキャッシュサーバに代わって、当該他のキャッシュサーバが最後に配信したストリームデータの続きを配信するようデータ配信部へ指示する配信不能検知部をさらに備える。

【0020】第11の発明は、第5の発明に従属する発明であって、第iのキャッシュサーバは、データ配信部から、重複して同じストリームデータを記憶する第（i+1）のキャッシュサーバおよび第（i-1）のキャッシュサーバに、端末装置へ現在配信している配信情報を、一定時間間隔で通知しており、第（i+1）のキャッシュサーバおよび第（i-1）のキャッシュサーバから通知される各々の配信情報によって、いずれか一方または

40 双方のキャッシュサーバが配信不能になったことを検知した場合、ストリームデータが途切れないよう定められた時間以内に、当該配信不能になったキャッシュサーバに代わって、当該配信不能になったキャッシュサーバが最後に配信したストリームデータの続きを配信するようデータ配信部へ指示する配信不能検知部をさらに備える。

【0021】第12および第13の発明は、それぞれ第10および第11の発明に従属する発明であって、配信不能検知部の指示に従って、データ配信部が、配信不能となったキャッシュサーバに代わって端末装置へストリ

ームデータを配信している場合、キャッシュ制御部は、ストリームデータ記憶部に新たなストリームデータを記憶しないように制御することを特徴とする。

【0022】上記のように、第10～第13の発明によれば、1つのストリームデータを複数のキャッシュサーバに記憶させておき、さらに、相互にキャッシュサーバの配信状態を通知しながら運用する。これにより、上述した発明の効果に加え、障害発生時も端末装置への配信が途切れることなく継続させることができる。

【0023】第14および第15の発明は、それぞれ第10および第11の発明に従属する発明であって、配信不能となっていたキャッシュサーバが復旧した場合、復旧したキャッシュサーバは、重複して同じストリームデータを記憶する他のキャッシュサーバのデータ配信部へ復旧通知を出力し、配信不能時に、復旧したキャッシュサーバに代わって端末装置へストリームデータを配信していた他のキャッシュサーバは、復旧通知を受けて、ストリームデータ記憶部に記憶していたストリームデータを、復旧したキャッシュサーバのストリームデータ記憶部にコピーすることを特徴とする。

【0024】上記のように、第14および第15の発明によれば、1つのストリームデータを複数のキャッシュサーバに記憶させておき、相互にキャッシュサーバの配信状態を通知しながら運用することを利用して、障害が発生した後に復旧したキャッシュサーバに対し迅速にリカバリ処理を行うことが可能となる。

【0025】第16および第17の発明は、それぞれ第1および第5の発明に従属する発明であって、キャッシュサーバは、端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションと、それぞれのセッションに関してデータ配信部が最後に送信したパケット（ストリームデータの一単位）と、ストリームデータの配信ストリーム数とを、管理するセッション管理部と、セッション管理部で管理されている全てのセッションのパケットを、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知する配信不能検知部とをさらに備え、配信不能検知部は、通知が一定時間来ない場合、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないよう定められた時間以内に、当該配信不能になったキャッシュサーバに代わって、同じストリームデータを記憶するいずれか1台の他のキャッシュサーバが、当該配信不能になったキャッシュサーバで管理される全てのセッションに関して、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するようデータ配信部へ指示することを特徴とする。

【0026】第18および第19の発明は、それぞれ第1および第5の発明に従属する発明であって、キャッシュサーバは、端末装置へストリームデータを配信するためのネットワーク接続情報であるセッションを識別する

ためのセッション識別子と、それぞれのセッションに関してデータ配信部が最後に送信したパケットを識別するためのパケット識別子と、当該セッション識別子およびパケット識別子を格納するためのセッション管理テーブルと、当該セッション管理テーブルを当該セッション識別子の値によって2つに分類するためのテーブル境界値と、ストリームデータの配信ストリーム数とを、管理するセッション管理部と、セッション管理部で管理されているストリームデータを配信中の全てのセッションのセ

10 ッション識別子およびパケット識別子を、重複して同じストリームデータを記憶する複数の他のキャッシュサーバとの間で、定期的に通知し、当該他のキャッシュサーバから受け取る通知に従って、セッション管理テーブルに格納されているセッション識別子およびパケット識別子の値を更新する、配信不能検知部とをさらに備え、セッション管理部は、自己のデータ配信部が配信しているストリームデータのセッションに関しては、セッション識別子をテーブル境界値より小さい値に設定し、他のキャッシュサーバのデータ配信部が配信しているストリームデータのセッションに関しては、当該他のキャッシュサーバで管理されるセッション識別子の値を、テーブル境界値より大きい値に変換して設定し、データ配信部は、全てのセッションにおいて、セッション識別子の値がテーブル境界値より小さい場合は、パケットを端末装置へ送信すると共に、パケット識別子の値を更新することを繰り返し、配信不能検知部は、通知が一定時間来ない場合は、通知元のキャッシュサーバが配信不能になったと判断し、ストリームデータが途切れないよう定められた時間以内に、テーブル境界値をセッション識別子の最大値より大きい値に設定することを特徴とする。

【0027】第20および第21の発明は、それぞれ第1および第5の発明に従属する発明であって、キャッシュサーバは、自己が配信しているストリームデータの配信ストリーム数を管理するセッション管理部と、他のキャッシュサーバが配信不能になった場合、当該配信不能になった他のキャッシュサーバを検知し、当該他のキャッシュサーバが配信していた全てのストリームデータの配信を、当該他のキャッシュサーバに代わって継続することをデータ配信部へ指示する配信不能検知部とをさらに備え、配信不能検知部は、セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、同じストリームデータを記憶するキャッシュサーバのうち、配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するようデータ配信部へ指示し、あふれ通知を受け取ったさらに他のキャッシュサーバにおける配信不能検知部は、あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリーム

データの全てについて、あふれ通知を出力したキャッシュサーバが最後に配信したストリームデータの続きを配信するようデータ配信部へ指示することを特徴とする。

【0028】第22および第23の発明は、それぞれ第16および第17の発明に従属する発明であって、他のキャッシュサーバが配信不能になった場合であって、セッション管理部に問い合わせた結果、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合、配信不能検知部は、同じストリームデータを記憶するキャッシュサーバのうち、配信不能になった他のキャッシュサーバ以外のさらに他のキャッシュサーバへあふれ通知を出力すると共に、最大配信ストリーム性能を超える分のストリームデータの配信を停止するようデータ配信部へ指示し、あふれ通知を受け取ったさらに他のキャッシュサーバにおける配信不能検知部は、あふれ通知を出力したキャッシュサーバが配信を停止した最大配信ストリーム性能を超える分のストリームデータ全てについて、最後に受け取った通知から識別されたパケットの次のパケットから配信を継続するようデータ配信部へ指示することを特徴とする。

【0029】第24および第25の発明は、それぞれ第18および第19の発明に従属する発明であって、同じストリームデータを記憶するキャッシュサーバは、最大セッション情報登録数がMAXのセッション管理テーブルをそれぞれ有し、それぞれのセッション管理部は、自己のデータ配信部が配信しているストリームデータのセッションに関しては、セッション識別子をテーブル境界値より小さい値に設定し、他のキャッシュサーバのデータ配信部が配信しているストリームデータのセッションに関しては、当該他のキャッシュサーバのセッション管理テーブルにおいて、テーブル境界値より小さいセッション識別子の値I1～Inを値(MAX-I1-1)～(MAX-In-1)に変換して設定し、配信不能検知部は、他のキャッシュサーバから受け取る通知に従つて、当該通知に含まれるセッション識別子I1～Inを(MAX-I1-1)～(MAX-In-1)に変換した値と、パケット識別子P1～Pnの値とを、セッション管理テーブルの値として更新し、通知されるテーブル境界値がFに変更された場合、テーブル境界値を(MAX-F)に設定することを特徴とする。

【0030】第26および第27の発明は、それぞれ第18および第19の発明に従属する発明であって、キャッシュサーバAにおいて、キャッシュサーバAとキャッシュサーバBとの両方に記憶されるストリームデータ集合S1と、キャッシュサーバAとキャッシュサーバCとの両方に記憶されるストリームデータ集合S2とを、ストリームデータ記憶部に記憶する場合、キャッシュサーバAが、セッション管理テーブルT0、T1と、セッション管理テーブルT0、T1をセッション識別子の値によって2つに分類するためのテーブル境界値F0、F1

とを有し、セッション管理テーブルT0には、キャッシュサーバAが配信中のストリームデータ集合S1に含まれるストリームデータのセッションに関しては、セッション識別子がテーブル境界値F0より小さい値に設定され、キャッシュサーバBが配信中のストリームデータ集合S1に含まれるストリームデータのセッションに関しては、キャッシュサーバBのセッション管理テーブルで、セッション識別子がテーブル境界値F0より小さい値I1～Inを値(MAX-I1-1)～(MAX-In-1)に変換されて設定され、セッション管理テーブルT1には、キャッシュサーバAが配信中のストリームデータ集合S2に含まれるストリームデータのセッションに関しては、セッション識別子がテーブル境界値F1より小さい値に設定され、キャッシュサーバCが配信中のストリームデータ集合S2に含まれるストリームデータのセッションに関しては、キャッシュサーバCのセッション管理テーブルで、セッション識別子がテーブル境界値F1より小さい値J1～Jnを値(MAX-J1-1)～(MAX-Jn-1)に変換されて設定され、キャッシュサーバAにおいて、配信不能検知部は、セッション管理部に問い合わせて、キャッシュサーバB、Cと互いにストリームデータを配信中の全てのセッションについて、セッション識別子およびパケット識別子を通知し、キャッシュサーバBからの通知が一定時間来ない場合は、テーブル境界値F0を、キャッシュサーバCからの通知が一定時間来ない場合は、テーブル境界値F1を、MAX以上に設定し、ストリーム配信数がキャッシュサーバの最大配信ストリーム性能を超える場合であつて、キャッシュサーバBからの通知が来ない場合は、テーブル境界値F1を、キャッシュサーバCからの通知が来ない場合は、テーブル境界値F0を、キャッシュサーバのストリーム配信数がキャッシュサーバの最大配信ストリーム性能以下になる値に変更することを特徴とする。

【0031】上記のように、第16～第27の発明によれば、1つのストリームデータを複数のキャッシュサーバに記憶させると共に、隣り合うキャッシュサーバの配信状態をセッションおよびパケットに関する情報形態で保持しておき、さらに、相互にキャッシュサーバの配信状態を通知しながら運用する。これにより、上述した発明の効果に加え、障害発生時も端末装置への配信が途切れることなく継続させることができる。

【0032】第28および第29の発明は、それぞれ第14および第15の発明に従属する発明であって、キャッシュサーバは、配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバの配信復旧部は、キャッシュ制御部がストリームデータの復旧を完了した後、代わりに端末装置へストリームデータを配

信していたキャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータの続きを配信するようデータ配信部へ指示し、キャッシュサーバの配信復旧部は、復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示することを特徴とする。

【0033】第30および第31の発明は、それぞれ第16および第17の発明に従属する発明であって、キャッシュサーバは、配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバのキャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバのデータ配信部へ復旧通知を出力し、代わりに端末装置へストリームデータを配信していたキャッシュサーバのストリームデータ記憶部に記憶されているストリームデータを、復旧した他のキャッシュサーバのストリームデータ記憶部にコピーし、復旧した他のキャッシュサーバの配信復旧部は、キャッシュ制御部がストリームデータの復旧を完了した後、キャッシュサーバへ復旧通知を出力すると共に、当該キャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット識別子で識別されるパケットの次のパケットから配信を継続するようデータ配信部へ指示し、キャッシュサーバの配信復旧部は、復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示することを特徴とする。

【0034】第32および第33の発明は、それぞれ第24および第25の発明に従属する発明であって、キャッシュサーバは、配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバのキャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバのデータ配信部へ復旧通知を出力し、代わりに端末装置へストリームデータを配信していたキャッシュサーバのストリームデータ記憶部に記憶されているストリームデータを、復旧した他のキャッシュサーバのストリームデータ記憶部にコピーし、復旧した他のキャッシュサーバの配信復旧部は、キャッシュ制御部がストリームデータの復旧を完了した後、キャッシュサーバへ復旧通知を出力すると共に、テーブル境界値を配信不能になる前の値に設定し、キャッシュサーバの配信復旧部は、テーブル境界値を、復旧した他のキャッシュサーバが配信不能になる前の値に設定することを特徴とする。

【0035】第34および第35の発明は、それぞれ第28および第29の発明に従属する発明であって、配信

不能になった他のキャッシュサーバの代わりにストリームデータを配信していたことで、キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、キャッシュサーバの配信復旧部は、復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示した後、復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータの続きを配信するようデータ配信部へ指示し、さらに他のキャッシュサーバの配信復旧部は、復旧通知を出力してきたキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示することを特徴とする。

【0036】第36および第37の発明は、それぞれ第30および第31の発明に従属する発明であって、配信不能になった他のキャッシュサーバの代わりにストリームデータを配信していたことで、キャッシュサーバが、最大配信ストリーム性能を超える分のストリームデータの配信を行っていた場合、キャッシュサーバの配信復旧部は、復旧した他のキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示した後、復旧した他のキャッシュサーバ以外のさらに他のキャッシュサーバへ復旧通知を出力すると共に、当該さらに他のキャッシュサーバが代わりに配信していたストリームデータにおいて、最後に送信したパケット識別子で識別されるパケットの次のパケットから配信を継続するようデータ配信部へ指示し、さらに他のキャッシュサーバの配信復旧部は、復旧通知を出力してきたキャッシュサーバの代わりに配信していたストリームデータの配信を、停止するようデータ配信部へ指示することを特徴とする。

【0037】第38および第39の発明は、それぞれ第26および第27の発明に従属する発明であって、キャッシュサーバは、配信不能となっていた他のキャッシュサーバが復旧したことを検知する配信復旧部をさらに備え、配信不能となっていた他のキャッシュサーバが復旧した場合、当該復旧した他のキャッシュサーバのキャッシュ制御部は、同じストリームデータを記憶しているキャッシュサーバのデータ配信部へ復旧通知を出力し、代わりに端末装置へストリームデータを配信していたキャッシュサーバAのストリームデータ記憶部に記憶されているストリームデータを、復旧した他のキャッシュサーバのストリームデータ記憶部にコピーし、復旧した他のキャッシュサーバの配信復旧部は、キャッシュ制御部が復旧通知を行ってストリームデータの復旧を完了した後、テーブル境界値を配信不能になる前の値に変更し、キャッシュサーバAの配信復旧部は、復旧したキャッシュサーバから通知されるテーブル境界値が、復旧したキャッシュサーバが配信不能になる前の値に戻った場合、

テーブル境界値  $F_0$ ,  $F_1$  を復旧したキャッシュサーバが配信不能になる前の値に戻すことを特徴とする。

【0038】上記のように、第28～第39の発明によれば、1つのストリームデータを複数のキャッシュサーバに記憶させると共に、隣り合うキャッシュサーバの配信状態をセッションおよびパケットに関する情報形態で保持しておき、相互にキャッシュサーバの配信状態を通知しながら運用することを利用して、障害が発生した後に復旧したキャッシュサーバに対し迅速にリカバリ処理を行うことが可能となる。また、このリカバリ処理によって障害が起こる前の状態に戻すことで、再度キャッシュサーバが配信不能になってしまって、端末装置への配信が途切れることなく継続させることができる。

【0039】

【発明の実施の形態】以下、本発明の実施の形態を、図面を用いて説明する。図1は、本発明の一実施形態に係るクラスタサーバ装置を用いた配信システムの構成を示すブロック図である。図1において、本実施形態に係るクラスタサーバ装置10を用いた配信システムは、m個の端末装置41～4n（mは、2以上の整数）とクラスタサーバ装置10とが、ネットワーク51で接続されている。クラスタサーバ装置10は、負荷分散サーバ20と、n個のキャッシュサーバ101～10n（nは、2以上の整数）と、コンテンツサーバ30とで構成される。このキャッシュサーバ101～10nは、内部ネットワーク52によって負荷分散サーバ20に、内部ネットワーク53によってコンテンツサーバ30に、それぞれ接続されている。

【0040】キャッシュサーバ101～10nは、セッション管理部111～11nと、データ配信部121～12nと、配信不能検知部131～13nと、ストリームデータ管理部141～14nと、ストリームデータ記憶部151～15nと、要求転送部161～16nと、キャッシュ制御部171～17nと、配信復旧部181～18nとを、それぞれ備えている。また、コンテンツサーバ30は、ストリームデータ蓄積部31と、読み出し部32と、キャッシュ33と、配信部34とを備えている。さらに、負荷分散サーバ20は、クラスタ制御部21で構成される。なお、図1に示すように、キャッシュサーバ101～10nは、それぞれ同一の構成であるので、以下、各キャッシュサーバ101～10nが同一処理を行う部分の説明に関しては、キャッシュサーバ10x（xは、1～nのいずれか）の表記を行うこととする。

【0041】まず、本実施形態に係るクラスタサーバ装置10の各構成の概要を説明する。クラスタ制御部21は、端末装置41～4nから送信される要求を受け付け、セッション管理部111～11nおよびストリームデータ管理部141～14nの情報に基づき、後述する手法を用いて、負荷が分散されるようにいずれかのキャ

ッシュサーバ101～10nに要求を転送する。また、クラスタ制御部21は、キャッシュサーバ101～10nが送出するストリームデータを入力し、対応する端末装置41～4nへ配信する。

【0042】キャッシュサーバ101において、セッション管理部111は、キャッシュサーバ101が行っているストリームデータの配信状態を管理する。データ配信部121は、要求転送部161の指示に従って、ストリームデータをクラスタ制御部21に出力する。配信不能検知部131は、システム管理上前後するキャッシュサーバ101..nまたはキャッシュサーバ101..nが、配信不能になったことを検知する。なお、x=1の場合

「x-1」は「n」に置き換えられ、x=nの場合「x+1」は「1」に置き換えられる。以下、変数i, j, kに関しても同様に置き換えられるものとする。配信復旧部181は、隣り合うキャッシュサーバ101..nまたはキャッシュサーバ101..nが、配信不能の状態から復旧したことを検知する。ストリームデータ記憶部151は、例えばRAM（random access memory）で構成され、コンテンツサーバ30から送信されてくるストリームデータを一時的に記憶する。ストリームデータ管理部141は、ストリームデータ記憶部151に記憶されているストリームデータの情報を保持する。要求転送部161は、クラスタ制御部21から転送される要求のストリームデータが、ストリームデータ記憶部151に記憶されているか否かを判断する。そして、要求転送部161は、記憶されている場合にはストリームデータ記憶部151に記憶されているストリームデータを、記憶されていない場合にはキャッシュ制御部171に対しコンテンツサーバ30からの配信要求を行うことで入手したストリームデータを、クラスタ制御部21へ出力するようにデータ配信部121へ指示する。キャッシュ制御部171は、要求転送部161から出力される配信要求をコンテンツサーバ30へ出力する。また、キャッシュ制御部171は、配信要求に対してコンテンツサーバ30から配信されてくるストリームデータを、ストリームデータ記憶部151に記憶するよう制御すると共に、システム管理上で連続するキャッシュサーバ101..nのストリームデータ記憶部151..nにも記憶するよう制御する。

【0043】コンテンツサーバ30において、ストリームデータ蓄積部31は、例えば大容量のハードディスクで構成され、複数のストリームデータをブロック単位で格納している。読み出し部32は、配信部34の指示に基づいて、ストリームデータ蓄積部31からストリームデータを読み出す。キャッシュ33は、読み出し部32が読み出したストリームデータを一時的に格納する。配信部34は、キャッシュサーバ101～10nから送信されてくる配信要求に従って読み出し部32を制御し、また、キャッシュ33に読み出されているストリームデ

10  
10  
20  
20  
30  
30  
40  
40

ータを対応するキャッシュサーバ $10_1 \sim 10_n$ へ配信する。

【0044】上記構成によって、本実施形態に係るクラスタサーバ装置 $10$ は、キャッシュサーバの負荷分散が最適となるストリームデータの配信動作、キャッシュサーバのいずれか1台に障害が発生して端末装置へのデータ配信ができなくなった場合でも、他のキャッシュサーバが代わってデータ配信を継続するフェールセーフ動作、および障害が発生して端末装置へのデータ配信ができなくなっていたキャッシュサーバが復旧した場合のリカバリ動作、をそれぞれ実現する。次に、本実施形態に係るクラスタサーバ装置 $10$ が実現する上記動作を、それぞれ詳細に説明する。

【0045】(1)ストリームデータ配信動作

まず、キャッシュサーバの負荷分散が最適となるストリームデータの配信動作を、図2および図3をさらに参照して説明する。図2および図3は、本発明の一実施形態に係るクラスタサーバ装置 $10$ が行うストリームデータの配信動作を説明するフローチャートである。なお、以下の説明では、図1に示したシステム環境と同様に、クラスタサーバ装置 $10$ は、 $n$ 個のキャッシュサーバ $10_1 \sim 10_n$ で構成されており、 $m$ 個の端末装置 $4_1 \sim 4_m$ とのデータ配信を行うものとする。

【0046】図2を参照して、端末装置 $4_y$  ( $y$ は、 $1 \sim m$ のいずれか) からストリームデータの配信要求が出力されると、クラスタ制御部 $21$ がこの配信要求を受信する(ステップS101)。配信要求を受信すると、クラスタ制御部 $21$ は、各キャッシュサーバ $10_1 \sim 10_n$ に対して、現在配信中であるストリーム数と、要求されているストリームデータがストリームデータ記憶部 $15_1 \sim 15_n$ に記憶されているか否かとを、問い合わせする(ステップS102)。この問い合わせに対して、各キャッシュサーバ $10_1 \sim 10_n$ は、セッション管理部 $11_1 \sim 11_n$ の情報に基づいて現在配信しているストリーム数を、ストリームデータ管理部 $14_1 \sim 14_n$ の情報に基づいて記憶の有無を、クラスタ制御部 $21$ へそれぞれ回答する。この各ストリームデータ管理部 $14_1 \sim 14_n$ は、例えば、図4のような情報を保持している。図4において、ストリームデータ管理部 $14_1$ は、ストリームデータの番号または名称と、ストリームデータ記憶部 $15_1$ 内の記憶位置と、前回アクセスした時刻とを保持している。このような情報により、キャッシュサーバ $10_1$ は、どんなストリームデータがストリームデータ記憶部 $15_1$ のどこに記憶管理されているかを判断する。次に、クラスタ制御部 $21$ は、全てのキャッシュサーバ $10_1 \sim 10_n$ から受けた回答に基づいて、要求されたストリームデータがストリームデータ記憶部 $15_1$ に記憶されているキャッシュサーバ $10_k$ が、存在するか否かを判断する(ステップS103)。

【0047】まず、クラスタ制御部 $21$ が、要求された

ストリームデータが記憶されているキャッシュサーバは存在しないと判断した場合を(ステップS103, N $o$ )、図2を続けて参照して説明する。以下に説明するステップS104～S111の動作は、基本的にクラスタサーバ装置 $10$ が行う初期設定動作に該当する。上記ステップS103において存在しないと判断した場合、クラスタ制御部 $21$ は、全てのキャッシュサーバ $10_1 \sim 10_n$ の中から現在配信中のストリーム数が一番小さいキャッシュサーバ $10_k$  ( $k$ は、 $1 \sim n$ のいずれか) を選択する(ステップS104)。

【0048】次に、クラスタ制御部 $21$ は、選択したキャッシュサーバ $10_k$ が配信中のストリーム数が、値 $\{((n-1)/n) \times MAX\}$ 未満か否かを判定する(ステップS105)。なお、MAXは、1台のキャッシュサーバが配信可能な最大ストリーム数を表す。上記ステップS105において未満であると判定した場合、クラスタ制御部 $21$ は、選択したキャッシュサーバ $10_k$ に対して、端末装置 $4$ から受け付けた要求を転送する(ステップS106)。なお、セッション管理部 $11_k$ が一定間隔で全てのキャッシュサーバ $10_1 \sim 10_n$ と通信し合い、キャッシュサーバ $10_1 \sim 10_n$ の配信状態を保持する負荷状態管理部を備えている場合には、クラスタ制御部 $21$ の代わりに負荷状態管理部が、キャッシュサーバ $10_k$ を選択し、端末装置 $4$ から受け付けた要求を転送するようにしてもよい。一方、上記ステップS105において未満でないと判定した場合、クラスタ制御部 $21$ は、要求の受け付けを拒絶する通知を端末装置 $4$ へ発行し、この配信動作を終了する(ステップS111)。

【0049】クラスタ制御部 $21$ から要求を転送されたキャッシュサーバ $10_k$ は、ストリームデータ記憶部 $15_k$ 内に要求されたストリームデータが記憶されていないので、コンテンツサーバ $30$ に対して配信要求を出力する。ここで、キャッシュ制御部 $17_k$ は、この配信要求を、システム管理上で連続するキャッシュサーバ $10_{k+1} \sim 10_n$ のキャッシュ制御部 $17_{k+1} \sim 17_n$ からの要求としても出力する(ステップS107)。これに対し、キャッシュサーバ $10_k$ からの配信要求を受信したコンテンツサーバ $30$ は、該当するストリームデータをストリームデータ蓄積部 $31$ から読み出し、キャッシュサーバ $10_k$ およびキャッシュサーバ $10_{k+1} \sim 10_n$ の双方に配信する(ステップS108)。

【0050】ストリームデータが配信されてきたキャッシュサーバ $10_k$ は、ストリームデータ記憶部 $15_k$ に当該ストリームデータを記憶すると共に、ストリームデータ管理部 $14_k$ の情報を更新する。なお、キャッシュサーバ $10_k$ は、上記ステップS107において、1対多通信をするためのマルチキャストアドレスを送信元アドレスに変更した配信要求をキャッシュ制御部 $17_k$ から出力した場合には、コンテンツサーバ $30$ から送信さ

れてきたマルチアドレス宛のストリームデータを、新たに設けるキャッシングデータ管理部が受信してストリームデータ記憶部15<sub>i</sub>に記憶するように制御すればよい。そして、キャッシングサーバ10<sub>i</sub>のデータ配信部12<sub>i</sub>は、ストリームデータ記憶部15<sub>i</sub>に記憶されたストリームデータを、クラスタ制御部21を介して端末装置4<sub>i</sub>へ配信する(ステップS109)。さらに、キャッシングサーバ10<sub>i</sub>は、端末装置4<sub>i</sub>へ配信したデータ内容を一定間隔でキャッシングサーバ10<sub>i+1</sub>に通知する(ステップS110)。この通知される情報は、例えば、図5に示すようなものであり、どの端末装置4<sub>i</sub>に対してどのストリームデータのどの部分を配信したかが分かる情報である。なお、キャッシングサーバ10<sub>i</sub>が行う配信情報の通知は、端末装置への配信ごとに行っててもよいし、全ての端末装置への配信に関して一括して行っててもよい。一方、キャッシングサーバ10<sub>i+1</sub>は、要求していないストリームデータがコンテンツサーバ30から配信されてきたので、ストリームデータ記憶部15<sub>i+1</sub>に当該ストリームデータを記憶すると共に、ストリームデータ管理部14<sub>i+1</sub>の情報を更新する(ステップS109)。

【0051】以上のように、ステップS104～S110の処理を行うことで、要求に基づいて配信されたストリームデータが、システム上連携して管理する連続した2つのキャッシングサーバに、それぞれ記憶されることになる。

【0052】次に、クラスタ制御部21が、要求されたストリームデータが記憶されているキャッシングサーバが存在すると判断した場合を(ステップS103, Yes)、図3を参照して説明する。上記ステップS103において存在すると判断した場合、クラスタ制御部21は、要求されたストリームデータを記憶していると回答した2つ以上のキャッシングサーバ10<sub>i</sub>の中で、システム管理上の番号が一番小さいキャッシングサーバ10<sub>i</sub> (iは、xが取り得る値)を選択する(ステップS121)。ここで、ストリームデータは、上述のように(上記ステップS109において)連続する一対のキャッシングサーバ10<sub>i</sub>およびキャッシングサーバ10<sub>i+1</sub>に必ず記憶されているので、クラスタ制御部21は、この一対のうち、キャッシングサーバ10<sub>i</sub>に関して番号が一番小さいキャッシングサーバ10<sub>i</sub>を選択する。例えば、ストリームデータを記憶していると回答したキャッシングサーバが、キャッシングサーバ10<sub>2</sub>, 10<sub>3</sub>の対と、キャッシングサーバ10<sub>1</sub>, 10<sub>2</sub>の対との2つであった場合、クラスタ制御部21は、キャッシングサーバ10<sub>2</sub>を選択する。

【0053】続いて、クラスタ制御部21は、選択したキャッシングサーバ10<sub>i</sub>が現在配信中のストリームデータのうち、ストリームデータ記憶部15<sub>i</sub>とキャッシングサーバ10<sub>i+1</sub>のストリームデータ記憶部15<sub>i+1</sub>と

に、記憶されているストリームデータの配信ストリーム数の合計が、MAX未満であるか否かを判定する(ステップS122)。このステップS122においてMAX未満でないと判定した場合、クラスタ制御部21は、ストリームデータを記憶していると回答した残りのキャッシングサーバ10<sub>j</sub> (jは、xが取り得る値)を選択し(ステップS128)、再びステップS122の判定を行う。例えば、上述した例において、キャッシングサーバ10<sub>1</sub>, 10<sub>2</sub>の配信ストリーム数の合計がMAX未満でない場合、クラスタ制御部21は、キャッシングサーバ10<sub>1</sub>を選択する。なお、ストリームデータを記憶していると回答したキャッシングサーバ10<sub>i</sub>が残っていない場合は、図2のステップS104に移行して以降の処理を行う(ステップS127)。一方、ステップS122においてMAX未満であると判定した場合、クラスタ制御部21は、次にキャッシングサーバ10<sub>i</sub>が配信中のストリーム数が、値 $\{( (n-1) / n ) \times MAX\}$ 未満か否かを判定する(ステップS123)。

【0054】ところで、クラスタ制御部21がステップS122, S123で行う判定は、例えば、図6のような情報に基づいて行われる。図6において、「キャッシングサーバ」はキャッシングサーバ10<sub>1</sub>～10<sub>6</sub>の番号を示し、「キャッシングの有無」は要求されたストリームデータがキャッシングされているか否かを表し、「ストリーム配信数1」はキャッシングサーバ10<sub>1</sub>とキャッシングサーバ10<sub>2</sub>とに記憶されているストリームデータの合計配信数を示し、「ストリーム配信数2」はキャッシングサーバ10<sub>1</sub>とキャッシングサーバ10<sub>2</sub>とに記憶されているストリームデータの合計配信数を示している。図6では、キャッシングサーバ10<sub>1</sub>が配信中のストリーム数は、120+1000=1120であり、キャッシングサーバ10<sub>1</sub>およびキャッシングサーバ10<sub>2</sub>に記憶されているストリームデータの配信数は、それぞれ120および500であることを表している。従って、ステップS122の判定では、620 (=120+500)とMAXとを比較し、ステップS123の判定では、1120と値 $\{( (n+1) / n ) \times MAX\}$ とを比較する。

【0055】なお、いずれか1台のキャッシングサーバに障害が発生した場合の対応を考慮する必要がなければ、キャッシングサーバ10<sub>i</sub>が配信中のストリーム数が、値 $\{( (n-1) / n ) \times MAX\}$ 未満か否かの判定ではなく、キャッシングサーバ10<sub>i</sub>が現在配信中のストリームデータのうち、ストリームデータ記憶部15<sub>i</sub>とキャッシングサーバ10<sub>i+1</sub>のストリームデータ記憶部15<sub>i+1</sub>とに、記憶されているストリームデータの配信ストリーム数の合計が、MAX未満であるか否かを判定してもよい。この場合、それぞれのストリームデータ記憶部15<sub>i</sub>を効率よく使用することができる。

【0056】再び図3を参照して、ステップS123に

おいてMAX未満であると判定した場合、クラスタ制御部21は、選択したキャッシュサーバ10<sub>i</sub>に対して端末装置4<sub>i</sub>から受け付けた要求を転送する（ステップS124）。クラスタ制御部21から要求を転送されたキャッシュサーバ10<sub>i</sub>は、データ配信部12<sub>i</sub>によってストリームデータ記憶部15<sub>i</sub>に記憶されているストリームデータを読み出し、クラスタ制御部21を介して端末装置4<sub>i</sub>へ配信する（ステップS125）。さらに、キャッシュサーバ10<sub>i</sub>は、端末装置4<sub>i</sub>へ配信したデータ内容を一定間隔でキャッシュサーバ10<sub>i+1</sub>へ通知する（ステップS126）。

〔0057〕一方、ステップS123においてMAX未満でないと判定した場合、クラスタ制御部21は、キャッシュサーバ10<sub>1,11</sub>が配信中のストリーム数が、値{(n-1)/n}×MAX}未満か否かをさらに判定する(ステップS129)。ステップS129において値未満であると判定した場合、クラスタ制御部21は、キャッシュサーバ10<sub>1,11</sub>に対して、端末装置4<sub>1</sub>から受け付けた要求を転送する(ステップS130)。クラスタ制御部21から要求を転送されたキャッシュサーバ10<sub>1,11</sub>は、データ配信部12<sub>1,11</sub>によってストリームデータ記憶部15<sub>1,11</sub>に記憶されているストリームデータを読み出し、クラスタ制御部21を介して端末装置4<sub>1</sub>へ配信する(ステップS131)。さらに、キャッシュサーバ10<sub>1,11</sub>は、端末装置4<sub>1</sub>へ配信したデータ内容を一定間隔でキャッシュサーバ10<sub>1</sub>へ通知する(ステップS132)。なお、ステップS129においてMAX未満でないと判定された場合、図2のステップS104に移行して、以降の処理を行う。

### 【0058】 (2) フェールセーフ動作

次に、クラスタサーバ装置を構成するキャッシュサーバのいずれか1台に障害が発生して、端末装置へのデータ配信ができなくなつた場合でも、他のキャッシュサーバが代わつてデータ配信を継続する、フェールセーフ動作について説明する。上述したように、キャッシュサーバ10<sub>1</sub>は、端末装置4<sub>1</sub>、ヘストリームデータを配信しながら、配信したストリームデータの情報を、キャッシュサーバ10<sub>111</sub>またはキャッシュサーバ10<sub>111</sub>へ一定間隔で通知している（ステップS110<sub>1</sub>、S126<sub>1</sub>、S132<sub>1</sub>）。これに対し、キャッシュサーバ10<sub>111</sub>の配信不能検知部13<sub>111</sub>およびキャッシュサーバ10<sub>111</sub>の配信不能検知部13<sub>111</sub>は、それぞれキャッシュサーバ10<sub>1</sub>から一定間隔で通知されてくる配信情報を、モニタリングしている。

【0059】今、キャッシュサーバ10<sub>1</sub>に障害が発生し、端末装置4<sub>1</sub>へのデータ配信ができなくなった場合を考える。なお、キャッシュサーバ10<sub>1</sub>とキャッシュサーバ10<sub>2</sub>とで、ある1つのストリームデータを連携して管理している場合を説明する。この場合、キャッシュサーバ10<sub>1</sub>の配信不能検知部13<sub>1</sub>は、キャ

10 ッシュサーバ10, から一定間隔で通知されてくる配信情報が来なくなったことにより、キャッシングサーバ10, の配信不能（障害発生）を検知する。配信不能検知部13,,, において配信不能が検知されると、データ配信部12,,, は、ストリームデータが途切れないよう予め定められた時間以内に、前回（障害発生前）にキャッシングサーバ10, から通知されてきた配信情報に基づいて、キャッシングサーバ10, が最後に配信したストリームデータの続きをストリームデータ記憶部15,,, から読み出して、端末装置4, へ配信する。ここで、ストリームデータが途切れないよう予め定められた時間とは、最後に配信されたストリームデータが端末装置4, で再生し終わるまでの時間である。なお、キャッシングサーバ10,,, とキャッシングサーバ10, とで、ある1つのストリームデータを連携して管理している場合であって、データ配信を行っているキャッシングサーバ10, に障害が発生した場合には、上記フェールセーフ動作は、キャッシングサーバ10,,, によって行われる。

【0060】このように、キャッシングサーバ10<sub>1..1</sub>、またはキャッシングサーバ10<sub>1..1</sub>が継続してストリームデータの配信を行うと、キャッシング制御部17<sub>1..1</sub>またはキャッシング制御部17<sub>1..1</sub>は、それぞれキャッシングサーバ10<sub>1..1</sub>の復旧を検知するまで、ストリームデータ記憶部15<sub>1..1</sub>またはストリームデータ記憶部15<sub>1..1</sub>に新たなストリームデータが記憶されないように、クラスタ制御部21に対してフェールセーフ状態であることを通知する。このフェールセーフ状態通知を受け取ったクラスタ制御部21は、新たにストリームデータを記憶させるキャッシングサーバ10<sub>1..1</sub>を選択する際に（ステップS104、S121）、キャッシングサーバ10<sub>1..1</sub>またはキャッシングサーバ10<sub>1..1</sub>を選択しないように制御する。

### 【0061】(3) リカバリ動作

次に、障害が発生して端末装置へのデータ配信ができないくなっていたキャッシュサーバが復旧した時のリカバリ動作を、キャッシュサーバ10<sub>1</sub>に発生していた障害が復旧した場合を一例に挙げて、説明する。障害が復旧したキャッシュサーバ10<sub>1</sub>は、キャッシュサーバ10<sub>1</sub>のデータ配信部1.2およびキャッシュサーバ1

40 0<sub>...1</sub> のデータ配信部 12<sub>...1</sub> に対して、データ配信部 12<sub>...1</sub> から復旧通知をそれぞれ出力する。復旧通知を受け取ったキャッシュサーバ 10<sub>...1</sub> のデータ配信部 12<sub>...1</sub> は、ストリームデータ記憶部 15<sub>...1</sub> に記憶されているストリームデータのうち、元々キャッシュサーバ 10<sub>...1</sub> とキャッシュサーバ 10<sub>...1</sub> の両方に記憶されていたストリームデータを選び出して、キャッシュサーバ 10<sub>...1</sub> のストリームデータ記憶部 15<sub>...1</sub> にコピーする。また、復旧通知を受け取ったキャッシュサーバ 10<sub>...1</sub> のデータ配信部 12<sub>...1</sub> は、ストリームデータ記憶部 15<sub>...1</sub> に記憶されているストリームデータのうち、元々

キャッシュサーバ10<sub>1</sub>、とキャッシュサーバ10<sub>2</sub>..、との両方に記憶されていたストリームデータを選び出して、キャッシュサーバ10<sub>1</sub>、のストリームデータ記憶部15<sub>1</sub>、にコピーする。そして、キャッシュサーバ10<sub>1</sub>、へのストリームデータのコピーが完了すると、キャッシュサーバ10<sub>1</sub>、のキャッシュ制御部17<sub>1</sub>、およびキャッシュサーバ10<sub>1</sub>、のキャッシュ制御部17<sub>2</sub>、..、は、クラスタ制御部21に対してフェールセーフ状態解除の通知を出力する。

【0062】(具体例を用いた動作説明) 次に、フェールセーフ動作とリカバリ動作とについて、具体的な一例を挙げてさらに詳しく説明する。ここでは、クラスタサーバ装置10が、4つのキャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、を備えている場合を、一例に挙げて説明する。また、図7は、キャッシュサーバ10<sub>1</sub>、内のストリームデータ記憶部15<sub>1</sub>、に記憶される、ストリームデータ601~60pの構成の一例を示す図である。各ストリームデータ601~60p(pは、正の整数)は、映像や音楽等の時間的に連続するデータである。ストリームデータ601~60pは、送信単位である複数のパケット611~61q(qは、正の整数)で、それぞれ構成される。また、パケット611~61qは、ストリームデータにおけるパケットの順番を識別するためパケット識別子621~62qを、それぞれ有している。

【0063】<前提条件その1:キャッシュサーバに記憶されるストリームデータの状態>上述したように、クラスタサーバ装置10では、要求に基づいて配信されたストリームデータが、システム上連携して管理する連続した2つのキャッシュサーバに、それぞれ記憶される。従って、この例では図8に示すように、キャッシュサーバ10<sub>1</sub>、のストリームデータ記憶部15<sub>1</sub>、と、キャッシュサーバ10<sub>2</sub>、のストリームデータ記憶部15<sub>2</sub>、とに、同一のストリームデータ701がそれぞれ記憶されている。また、キャッシュサーバ10<sub>1</sub>、のストリームデータ記憶部15<sub>1</sub>、と、キャッシュサーバ10<sub>2</sub>、のストリームデータ記憶部15<sub>2</sub>、とに、同一のストリームデータ702がそれぞれ記憶されている。また、キャッシュサーバ10<sub>3</sub>、のストリームデータ記憶部15<sub>3</sub>、と、キャッシュサーバ10<sub>4</sub>、のストリームデータ記憶部15<sub>4</sub>、とに、同一のストリームデータ703がそれぞれ記憶されている。また、キャッシュサーバ10<sub>1</sub>、のストリームデータ記憶部15<sub>1</sub>、と、キャッシュサーバ10<sub>2</sub>、のストリームデータ記憶部15<sub>2</sub>、とに、同一のストリームデータ704がそれぞれ記憶されている。

【0064】<前提条件その2:ストリーム配信数>上述したように、クラスタ制御部21は、次の条件を満たすように、各キャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、へ配信要求を転送する。なお、1台のキャッシュサーバが配信可能な最大ストリーム数MAXが、「8」であると仮定する。

【条件1】各キャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、のストリーム配信数は、6 (= (n-1) / n × MAX = 3 / 4 × 8) 以下となる。

【条件2】システム上連携して管理する連続した2つのキャッシュサーバに記憶されているデータ配信中のストリーム数は、8以下 (=MAX) となる。

【0065】<前提条件その3:セッション管理>各キャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、のセッション管理部11<sub>1</sub>、~11<sub>4</sub>、が行うセッション管理の手法について、図9

10 および図10を用いて説明する。セッション管理部11<sub>1</sub>、~11<sub>4</sub>、は、キャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、のストリーム配信数を、セッション管理テーブル53<sub>1</sub>、~53<sub>4</sub>、の状態で保持する。ここで、各キャッシュサーバ10<sub>1</sub>、~10<sub>4</sub>、は、2つずつセッション管理テーブルを保持する。すなわち、キャッシュサーバ10<sub>1</sub>、は、キャッシュサーバ10<sub>1</sub>、と共に記憶するストリームデータ701に関するセッションを、セッション管理テーブル53<sub>1</sub>、を使用して管理することでキャッシュサーバ10<sub>1</sub>、の配信状態を保持し、キャッシュサーバ10<sub>4</sub>、と共に記憶するストリームデータ704に関するセッションを、セッション管理テーブル53<sub>4</sub>、を使用して管理することでキャッシュサーバ10<sub>4</sub>、の配信状態を保持する。キャッシュサーバ10<sub>2</sub>、は、キャッシュサーバ10<sub>2</sub>、と共に記憶するストリームデータ702に関するセッションを、セッション管理テーブル53<sub>2</sub>、を使用して管理することでキャッシュサーバ10<sub>2</sub>、の配信状態を保持し、キャッシュサーバ10<sub>3</sub>、と共に記憶するストリームデータ701に関するセッションを、セッション管理テーブル53<sub>3</sub>、を使用して管理することでキャッシュサーバ10<sub>3</sub>、の配信状態を保持する。キャッシュサーバ10<sub>1</sub>、は、キャッシュサーバ10<sub>1</sub>、と共に記憶するストリームデータ703に関するセッションを、セッション管理テーブル53<sub>1</sub>、を使用して管理することでキャッシュサーバ10<sub>1</sub>、の配信状態を保持し、キャッシュサーバ10<sub>2</sub>、と共に記憶するストリームデータ702に関するセッションを、セッション管理テーブル53<sub>2</sub>、を使用して管理することでキャッシュサーバ10<sub>2</sub>、の配信状態を保持する。キャッシュサーバ10<sub>3</sub>、は、キャッシュサーバ10<sub>3</sub>、と共に記憶するストリームデータ704に関するセッションを、セッション管理テーブル53<sub>3</sub>、を使用して管理することでキャッシュサーバ10<sub>3</sub>、の配信状態を保持する。キャッシュサーバ10<sub>4</sub>、は、キャッシュサーバ10<sub>4</sub>、と共に記憶するストリームデータ701に関するセッションを、セッション管理テーブル53<sub>4</sub>、を使用して管理することでキャッシュサーバ10<sub>4</sub>、の配信状態を保持する。

【0066】例えば、キャッシュサーバ10<sub>1</sub>、とキャッシュサーバ10<sub>2</sub>、とで共通して記憶しているストリームデータ701に対して、ストリーム配信要求がキャッシュサーバ10<sub>1</sub>、またはキャッシュサーバ10<sub>2</sub>、に来た場合、図9中のセッション管理テーブル53<sub>1</sub>、およびセッ

ション管理テーブル 53<sub>1</sub> に、要求に関する情報をそれぞれ登録する。そして、ストリーム配信中は、セッション管理テーブル 53<sub>1</sub> およびセッション管理テーブル 53<sub>2</sub> にそれぞれ登録されている双方の情報を、その都度更新する。以下、図 10 を用いてさらに詳しく説明する。

【0067】各キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> は、端末装置 4<sub>1</sub> ～ 4<sub>4</sub> へストリームを配信するために、端末装置 4<sub>1</sub> ～ 4<sub>4</sub> およびクラスタ制御部 21 と、ネットワーク 51, 52 で接続されている。今、1つのストリームを配信するためのこれらのネットワークの接続状態を、セッションと呼ぶ。図 10 は、図 9 で示したセッション管理テーブル 53<sub>1</sub> およびセッション管理テーブル 53<sub>2</sub> の詳細を示す図である。セッション管理テーブル 53<sub>1</sub> およびセッション管理テーブル 53<sub>2</sub> には、セッションを識別するためのセッション識別子と、そのセッションで最後に送出したパケットを示すパケット識別子とが、それぞれ登録されている。以下の実施例では、各セッション管理テーブル 53<sub>1</sub> ～ 53<sub>4</sub> のセッション登録数の最大値を、MAX (= 8) と定めることにする。

【0068】セッション管理テーブル 53<sub>1</sub> とセッション管理テーブル 53<sub>2</sub> とでは、セッションの識別が上下逆に（テーブル領域において対称的な位置で）行われる。つまり、セッション管理テーブル 53<sub>1</sub> においてセッション識別子 I (0 ≤ I < MAX) で識別されるセッションは、セッション管理テーブル 53<sub>2</sub> においては、セッション識別子 (MAX - I - 1) で識別される。例えば、セッション管理テーブル 53<sub>1</sub> において、セッション識別子「0」、「1」で識別されるセッションのパケット識別子「100」、「510」は、セッション管理テーブル 53<sub>2</sub> において、セッション識別子「7」、「6」としてそれぞれ登録される。テーブル境界値 54<sub>1</sub> およびテーブル境界値 54<sub>2</sub> は、セッション管理テーブル 53<sub>1</sub> およびセッション管理テーブル 53<sub>2</sub> に登録されている情報を2つに分割するための、セッション識別子の値である。上述したように、セッション管理テーブル 53<sub>1</sub> とセッション管理テーブル 53<sub>2</sub> とでは、セッションの登録が上下逆に行われている。このため、一方のセッション管理テーブルのテーブル境界値が、F (0 ≤ F < MAX) に設定される場合、他方のセッション管理テーブルのテーブル境界値は、(MAX - F) に設定される。図 10 では、セッション管理テーブル 53<sub>1</sub> のテーブル境界値 54<sub>1</sub> に「3」が設定されているため、セッション管理テーブル 53<sub>2</sub> のテーブル境界値 54<sub>2</sub> には、「5」が設定されている。

【0069】キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> のデータ配信部 12<sub>1</sub> ～ 12<sub>4</sub> は、テーブル境界値 F より小さいセッション識別子 I を持つセッションに関してのみ、ストリームデータを端末装置へ送信する。具体的には、キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> は、下記の（第1ステップ

）および（第2ステップ）の動作を、繰り返す。（第1ステップ）セッション識別子 I とテーブル境界値 F を比較し、I < F であれば、セッション管理テーブル中のパケット識別子で示されるパケットの次のパケットを、ストリームデータ記憶部から取り出して端末装置へ送出する。（第2ステップ）セッション管理テーブル中のセッション識別子 I に関するパケット識別子の値を、上記第1ステップで送出したパケットのパケット識別子の値に更新する。

【0070】例えば、図 10 において、キャッシュサーバ 10<sub>1</sub> のセッション管理テーブル 53<sub>2</sub> におけるセッション識別子「0」～「2」で識別されるセッション（同図中、太線部分）に関しては、キャッシュサーバ 10<sub>1</sub> のデータ配信部 12<sub>1</sub> が、ストリームデータの送信およびパケット識別子の値の更新を行う。同様に、キャッシュサーバ 10<sub>2</sub> のセッション管理テーブル 53<sub>1</sub> におけるセッション識別子「0」～「4」で識別されるセッション（同図中、太線部分）に関しては、キャッシュサーバ 10<sub>2</sub> のデータ配信部 12<sub>2</sub> が、ストリームデータの送信およびパケット識別子の値の更新を行う。

【0071】<前提条件その4：配信不能検知部の動き>各セッション管理テーブルにおいて、テーブル境界値 F より大きいセッション識別子 I を持つセッションは、別のキャッシュサーバがストリーム配信中のセッションである。また、上述したように、キャッシュサーバ 10<sub>1</sub> の配信不能検知部 13<sub>1</sub> とキャッシュサーバ 10<sub>1,1</sub> の配信不能検知部 13<sub>1,1</sub> とは、互いに通信を行う。この配信不能検知部の動きを、図 11 および 12 を用いて説明する。図 11 は、n = 4 の場合における配信不能検知部 13<sub>1</sub> ～ 13<sub>4</sub> の動作の一例を示す図である。配信不能検知部 13<sub>1</sub> ～ 13<sub>4</sub> は、隣り合うキャッシュサーバと定期的に通信を行う。例えば、配信不能検知部 13<sub>1</sub> は、配信不能検知部 13<sub>2</sub> および配信不能検知部 13<sub>3</sub> と、定期的に通信を行う。そして、配信不能検知部 13<sub>3</sub> は、他のキャッシュサーバ 10<sub>2</sub> またはキャッシュサーバ 10<sub>4</sub> が配信不能になった場合、当該他のキャッシュサーバ 10<sub>2</sub> またはキャッシュサーバ 10<sub>4</sub> の代わりにストリームデータの配信を継続して開始するよう、データ配信部 12<sub>1</sub> へ指示する。

【0072】図 12 は、図 11 に示した配信不能検知部 13<sub>1</sub> と配信不能検知部 13<sub>2</sub> とにおいて、双方間で行われる通知と、その通知を受け取った時の動作とを、模式的に表した図である。セッション管理テーブル 53<sub>1</sub> およびセッション管理テーブル 53<sub>2</sub> は、キャッシュサーバ 10<sub>1</sub> およびキャッシュサーバ 10<sub>2</sub> がそれぞれ保持するセッション管理テーブルである。配信不能検知部 13<sub>1</sub> と配信不能検知部 13<sub>2</sub> とは、テーブル境界値と、前回の通知からパケット識別子の内容が更新されたセッションのセッション識別子およびパケット識別子の値と、定期的に通知する。

40

【0073】 例えば、前回から今回の通知の間に、キャッシュサーバ10<sub>1</sub>において、セッション識別子「0」、「2」で識別されるセッションのパケット識別子が、それぞれ「100」から「101」へ、「200」から「201」へ更新された場合、配信不能検知部13<sub>1</sub>は、自己のテーブル境界値と更新されたパケット識別子の値とを、配信不能検知部13<sub>2</sub>へ通知する。また、前回から今回の通知の間に、キャッシュサーバ10<sub>1</sub>において、セッション識別子「4」で識別されるセッションのパケット識別子が、「220」から「221」へ更新された場合、配信不能検知部13<sub>2</sub>は、自己のテーブル境界値と更新されたパケット識別子の値とを、配信不能検知部13<sub>1</sub>へ通知する。なお、テーブル境界値は、毎回通知されるが、通知される値Fまたは(MAX-F)が現在の値(MAX-F)またはFと異なる場合のみ、テーブル境界値の変更を行う。図12の例では、テーブル境界値は変更されない。今回の通知によって、配信不能検知部13<sub>1</sub>は、セッション識別子「4」のパケット識別子が「221」に更新されたことを検知する。セッション管理テーブル53<sub>1</sub>上のセッション識別子1に対応するパケット識別子は、セッション管理テーブル53<sub>2</sub>上では、セッション識別子「3(=MAX-I-1)」に格納されているため、配信不能検知部13<sub>2</sub>は、セッション識別子「3」のパケット識別子の値を「221」に更新する。同様に、今回の通知によって、配信不能検知部13<sub>2</sub>は、セッション識別子「0」、「2」のパケット識別子が「101」、「201」にそれぞれ更新されたことを検知する。そして、配信不能検知部13<sub>2</sub>は、セッション識別子「7」、「5」のパケット識別子の値を、それぞれ「101」、「201」に更新する。

【0074】 <基本動作その1：隣接するキャッシュサーバが障害を起こした場合> 図13に、1台のキャッシュサーバ10<sub>1</sub>が障害により配信不能になった場合の動作を示す。キャッシュサーバ10<sub>1</sub>が障害により配信不能になった場合、配信不能検知部13<sub>1</sub>からの通知が配信不能検知部13<sub>2</sub>へ来なくなる。一定時間通知を受け取らなかった配信不能検知部13<sub>2</sub>は、テーブル境界値54<sub>2</sub>をMAXに変更する。図13では、MAX=8であるので、テーブル境界値54<sub>2</sub>は「5」から「8」に変更される。その結果、キャッシュサーバ10<sub>2</sub>のデータ配信部12<sub>2</sub>は、セッション管理テーブル53<sub>2</sub>上のセッション識別子「5」～「7」で識別されるセッションが、テーブル境界値F=8より小さいセッション識別子になったと判断し、パケット識別子で示されるパケットの次のパケットから、ストリームデータの配信を継続して開始する。

【0075】 <基本動作その2：セッションのあふれ(テーブル境界値移動)> 図14に、キャッシュサーバ10<sub>1</sub>の最大配信数がオーバーする場合に、テーブル境

界値を変更させる(テーブル境界を移動させる)動作を示す。下記のいずれかの条件を満たす場合、キャッシュサーバ10<sub>1</sub>は、テーブル境界値を変更する。

(1) キャッシュサーバ10<sub>1</sub>の障害を検知して、キャッシュサーバ10<sub>1</sub>の代わりにキャッシュサーバ10<sub>2</sub>がストリームデータの配信を開始し、その結果、最大ストリーム配信性能を超えてしまった場合。

(2) キャッシュサーバ10<sub>1</sub>のテーブル境界値が変更されたことを検知して、キャッシュサーバ10<sub>1</sub>の代わりにキャッシュサーバ10<sub>2</sub>がストリームデータの配信を開始し、その結果、最大ストリーム配信性能を超えてしまった場合。

【0076】 図14では、まず、テーブル境界値54<sub>2</sub>が「3」から「1」に変更される。その結果、キャッシュサーバ10<sub>2</sub>のデータ配信部12<sub>2</sub>は、セッション管理テーブル53<sub>2</sub>上でセッション識別子「1」、「2」で識別されるセッションが、テーブル境界値F=1以上になったと判断し、ストリームデータの配信を停止する。その後、配信不能検知部13<sub>2</sub>から配信不能検知部

20 13<sub>1</sub>へ、テーブル境界値F=1が通知される。配信不能検知部13<sub>1</sub>は、通知によってテーブル境界値54<sub>1</sub>が変更されたことを検知し、テーブル境界値54<sub>1</sub>の値を(MAX-(通知されたテーブル境界値54<sub>2</sub>))に変更する。図14では、MAX=8であるので、テーブル境界値54<sub>1</sub>は「5」から「7(=8-1)」に変更される。その結果、キャッシュサーバ10<sub>2</sub>のデータ配信部12<sub>2</sub>は、セッション管理テーブル53<sub>2</sub>上のセッション識別子「5」、「6」で識別されるセッションが、テーブル境界値F=7より小さいセッション識別子になったと判断し、パケット識別子で示されるパケットの次のパケットから、ストリームデータの配信を継続して開始する。

【0077】 <基本動作その3：隣接するキャッシュサーバが復旧した場合> 上述したように、復旧したキャッシュサーバ10<sub>1</sub>のキャッシュ制御部17<sub>1</sub>は、隣接するキャッシュサーバ10<sub>2</sub>、およびキャッシュサーバ10<sub>1</sub>のデータ配信部12<sub>1</sub>、およびデータ配信部12<sub>2</sub>へ、復旧通知を出力する。そして、復旧したキャッシュサーバ10<sub>1</sub>は、配信されてくるストリームデータをストリームデータ記憶部15<sub>1</sub>に記憶させることで、ストリームデータの復旧を行う。図15は、ストリームデータの復旧完了後、障害を起こして配信不能になっていたキャッシュサーバ10<sub>1</sub>が、ストリームデータの配信を復旧する場合の動作を示す。

【0078】 キャッシュサーバ10<sub>1</sub>は、ストリームデータが復旧される前は、テーブル境界値54<sub>1</sub>が「0」であり、ストリームデータの配信を行っていない。配信復旧部18<sub>1</sub>は、キャッシュサーバ10<sub>1</sub>のストリームデータの復旧完了を待って、テーブル境界値54<sub>1</sub>を、障害を起こす前の値である「3」に変更する。その結

果、キャッシュサーバ10<sub>1</sub>のデータ配信部12<sub>1</sub>は、セッション管理テーブル53<sub>1</sub>上のセッション識別子「0」～「2」で識別されるセッションが、テーブル境界値F=3より小さいセッション識別子になったと判断し、パケット識別子で示されるパケットの次のパケットから、ストリームデータの配信を開始する。次に、配信不能検知部13<sub>1</sub>は、配信不能検知部13<sub>1</sub>ヘテーブル境界値54<sub>1</sub>を通知する。配信不能検知部13<sub>1</sub>は、通知によってテーブル境界値54<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>が障害を起こす前の値である「3」に戻ったことを検知し、テーブル境界値54<sub>1</sub>を(MAX-(通知されたテーブル境界値54<sub>1</sub>))に変更する。図15では、MAX=8であるので、テーブル境界値54<sub>1</sub>は「8」から「5」に戻される。その結果、キャッシュサーバ10<sub>1</sub>のデータ配信部12<sub>1</sub>は、セッション管理テーブル53<sub>1</sub>上のセッション識別子「5」～「7」で識別されるセッションが、テーブル境界値F=5より大きいセッション識別子になったと判断し、ストリームデータの配信を停止する。

【0079】<基本動作その4：セッション戻す（テーブル境界値移動）>基本的な動作は、上述した<基本動作その2>と同様である。配信復旧部18<sub>1</sub>は、テーブル境界値54<sub>1</sub>をキャッシュサーバ10<sub>1</sub>が障害を起こす前の値に設定すると共に、キャッシュサーバ10<sub>1</sub>が保持するテーブル境界値54<sub>1</sub>ではないもう一方のテーブル境界値（キャッシュサーバ10<sub>1</sub>とキャッシュサーバ10<sub>2</sub>とに記憶されるストリームデータのセッションを管理するセッション管理テーブルを、分割するテーブル境界値）を元の値に戻す。

【0080】（フェールセーフ動作例）上述した前提条件その1～4および基本動作その1～4を踏まえた上で、本発明の一実施形態に係るクラスタサーバ装置が行うフェールセーフ動作の具体的な一例を、図16および17に示す。図16は、最大配信ストリーム性能が「8」である4台のキャッシュサーバ10<sub>1</sub>～10<sub>4</sub>を用いて、ストリーミング配信が正常に行われている場合の、各キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>が管理するセッション管理テーブル53<sub>1</sub>～53<sub>4</sub>およびテーブル境界値54<sub>1</sub>～54<sub>4</sub>の状態の一例を表す図である。なお、図16において、各テーブル境界値54<sub>1</sub>～54<sub>4</sub>は「4」である。また、各キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>の配信ストリーム数（図中太線のセッションの合計数）は、全て「6」である。図17は、図16において、キャッシュサーバ10<sub>1</sub>が障害により配信不能になった場合の、各キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>が管理するセッション管理テーブル53<sub>1</sub>～53<sub>4</sub>およびテーブル境界値54<sub>1</sub>～54<sub>4</sub>の状態の一例を表す図である。

【0081】図16の状態において、キャッシュサーバ10<sub>1</sub>が障害によって配信不能になった場合は、次のよ

うな処理が行われる。まず、キャッシュサーバ10<sub>1</sub>の配信不能検知部13<sub>1</sub>およびキャッシュサーバ10<sub>1</sub>の配信不能検知部13<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>が配信不能になったことを検知し、テーブル境界値54<sub>1</sub>およびテーブル境界値54<sub>1</sub>を「8」に設定する。これに伴って、キャッシュサーバ10<sub>1</sub>の最大配信数が「10」になるので、最大ストリーム配信性能「8」を超える2つのセッションをキャッシュサーバ10<sub>1</sub>に配信させるため、配信不能検知部13<sub>1</sub>は、テーブル境界値54<sub>1</sub>を「2」に設定する。同様にして、キャッシュサーバ10<sub>1</sub>のテーブル境界値54<sub>1</sub>が「6」に設定されるが、キャッシュサーバ10<sub>1</sub>の最大配信数が最大ストリーム配信性能「8」を超えないで、問題ない。また、キャッシュサーバ10<sub>1</sub>のテーブル境界値54<sub>1</sub>が「8」に設定されるが、キャッシュサーバ10<sub>1</sub>の最大配信数が最大ストリーム配信性能「8」を超えないで、問題ない。これにより、キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>は、全て最大ストリーム配信性能「8」で各セッションのストリーム配信を行うこととなる。従って、キャッシュサーバ10<sub>1</sub>が障害を起こしても、全てのセッションに関して、ストリームデータを途切れることなく配信することができる。

【0082】（リカバリ動作例）図17の状態において、キャッシュサーバ10<sub>1</sub>が復旧した場合には、下記の第1～第6ステップに基づいて、ストリームデータのリカバリ動作が行われる。

（第1ステップ）キャッシュサーバ10<sub>1</sub>のキャッシュ制御部17<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>のストリームデータ記憶部15<sub>1</sub>～15<sub>4</sub>に記憶されている全てのストリームデータを、自己のストリームデータ記憶部15<sub>1</sub>にコピーする。

（第2ステップ）配信不能検知部13<sub>1</sub>は、キャッシュサーバ10<sub>1</sub>～10<sub>4</sub>の配信不能検知部13<sub>1</sub>～13<sub>4</sub>から通知される情報に基づいて、セッション管理テーブル53<sub>1</sub>～53<sub>4</sub>にセッション識別子およびパケット識別子を登録する。

（第3ステップ）キャッシュサーバ10<sub>1</sub>の配信復旧部18<sub>1</sub>が、テーブル境界値54<sub>1</sub>～54<sub>4</sub>を「4」に設定する。

（第4ステップ）キャッシュサーバ10<sub>1</sub>の配信復旧部18<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>からの通知に基づいてテーブル境界値54<sub>1</sub>が元の値に戻ったことを検知し、テーブル境界値54<sub>1</sub>～54<sub>4</sub>を「4」に設定する。

（第5ステップ）キャッシュサーバ10<sub>1</sub>の配信復旧部18<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>からの通知に基づいてテーブル境界値54<sub>1</sub>が元の値に戻ったことを検知し、テーブル境界値54<sub>1</sub>を「4」に設定する。

（第6ステップ）キャッシュサーバ10<sub>1</sub>の配信復旧部18<sub>1</sub>が、キャッシュサーバ10<sub>1</sub>からの通知に基づい

てテーブル境界値 54<sub>1</sub> が元の値に戻ったことを検知し、テーブル境界値 54<sub>1</sub> を「4」に設定する。

以上の処理により、各キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> の状態は、図 16 に示す状態に戻る。従って、再度いずれかのキャッシュサーバが配信不能になってしまっても、再度フェールセーフ動作を行うことが可能となる。

【0083】以上のように、本発明の一実施形態に係るクラスタサーバ装置によれば、1つのストリームデータを複数のキャッシュサーバに記憶しておく。これにより、複数のキャッシュサーバからストリームデータを直接配信することが可能となるので、配信性能の飛躍的向上を図ることができる。また、記憶されているストリームデータに応じてキャッシュサーバの負荷を均等に分散させることで、記憶部（キャッシュ）の効率的な運用、かつ、より多くの端末装置へのデータ配信が可能になる。また、常に複数のキャッシュサーバにストリームデータを記憶させておき、相互にキャッシュサーバの配信状態を通知しながら運用することで、障害発生時も端末装置への配信が途切れることなく継続させることができる。さらに、常に複数のキャッシュサーバにストリームデータを記憶させておるので、障害が発生した後に復旧したキャッシュサーバに対して迅速にリカバリ処理を行うことが可能となる。また、このリカバリ処理によって障害が起こる前の状態に戻すことで、再度キャッシュサーバが配信不能になってしまっても、端末装置への配信が途切れることなく継続させることができる。

【0084】なお、上記実施形態では、それぞれのキャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> がセッション管理部 11<sub>1</sub> ～ 11<sub>4</sub> とストリームデータ管理部 14<sub>1</sub> ～ 14<sub>4</sub> とを備えており、クラスタ制御部 21 からの問い合わせに対して配信中のストリーム数および記憶の有無を回答する方法を一例に挙げて説明した。しかし、この他にもクラスタサーバ装置 10 全体で1つのデータベースを保持し、全てのキャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> が個々にデータベースを更新することにより、クラスタ制御部 21 がデータベースを毎回参照することで、配信中のストリーム数および記憶の有無を把握できるようする方法でもよい。また、上記実施形態では、負荷分散サーバ 20 とキャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> とを別構成として説明したが、いずれかのキャッシュサーバ 10<sub>1</sub> に負荷分散サーバ 20 の機能を持たせるような実装形態や、全てのキャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> に負荷分散サーバ 20 の機能を持たせ、要求を受信したキャッシュサーバ 10<sub>1</sub> が負荷分散サーバとして動作するような形態でもよい。さらに、上記実施形態では、1台のキャッシュサーバが配信可能な最大ストリーム数を MAX とし、これを基準にして配信可能なストリーム数を決定した。しかし、複数のキャッシュサーバで性能が異なり最大ストリーム数を一意に決められない場合には、一番性能の低いキャッシュサーバに合わせて設定する、または

各キャッシュサーバの平均値に合わせて設定する、あるいは予め1台のキャッシュサーバが配信可能な最大ストリーム数に余裕を持たせて少なめに設定しておく等の方法で決定してもよい。

【図面の簡単な説明】

【図 1】本発明の一実施形態に係るクラスタサーバ装置を用いた配信システムの構成を示すブロック図である。

【図 2】本発明の一実施形態に係るクラスタサーバ装置 10 が行うストリームデータの配信動作を説明するフローチャートである。

【図 3】本発明の一実施形態に係るクラスタサーバ装置 10 が行うストリームデータの配信動作を説明するフローチャートである。

【図 4】ストリームデータ管理部 14<sub>1</sub> ～ 14<sub>4</sub> が格納している情報の一例を示す図である。

【図 5】データ配信部 12<sub>1</sub> ～ 12<sub>4</sub> が通知する配信情報の一例を示す図である。

【図 6】クラスタ制御部 21 が行う判定動作に用いる情報の一例を示す図である。

【図 7】ストリームデータ記憶部 15<sub>1</sub> ～ 15<sub>4</sub> が保持するストリームデータの構成の一例を示す図である。

【図 8】各キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> のストリームデータ記憶部 15<sub>1</sub> ～ 15<sub>4</sub> が記憶するストリームデータの状態の一例を示す図である。

【図 9】各キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> のセッション管理部 11<sub>1</sub> ～ 11<sub>4</sub> が保持するセッション管理テーブル 53<sub>1</sub> ～ 53<sub>4</sub> の状態の一例を示す図である。

【図 10】図 9 のセッション管理テーブル 53<sub>1</sub> ～ 53<sub>4</sub> を詳細に表した図である。

【図 11】各キャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> の配信不能検知部 13<sub>1</sub> ～ 13<sub>4</sub> が、どのキャッシュサーバの配信不能検知部と通知を行うかを示す図である。

【図 12】図 11 の配信不能検知部 13<sub>1</sub> ～ 13<sub>4</sub> が互いに通知する内容を表した図である。

【図 13】一方のキャッシュサーバ 10<sub>1</sub> が配信不能になった場合の、他のキャッシュサーバ 10<sub>2</sub> の配信不能検知部 13<sub>2</sub> の動作を表した図である。

【図 14】一方のキャッシュサーバ 10<sub>1</sub> のストリーム配信の一部があふれる場合の、他のキャッシュサーバ 10<sub>2</sub> の配信不能検知部 13<sub>2</sub> の動作を表した図である。

【図 15】配信不能になっていたキャッシュサーバ 10<sub>1</sub> が復旧する際の、配信復旧部 18<sub>1</sub> および配信不能検知部 13<sub>1</sub> の動作を表した図である。

【図 16】4台のキャッシュサーバ 10<sub>1</sub> ～ 10<sub>4</sub> で配信を行う場合の、セッション管理テーブル 53<sub>1</sub> ～ 53<sub>4</sub> およびテーブル境界値 54<sub>1</sub> ～ 54<sub>4</sub> の状態を表す図である。

【図 17】図 16において、1台のキャッシュサーバ 10<sub>1</sub> が配信不能になった場合の、セッション管理テーブル 53<sub>1</sub> ～ 53<sub>4</sub> およびテーブル境界値 54<sub>1</sub> ～ 54<sub>4</sub>

の状態を表す図である。

【図 18】従来のクラスタサーバ装置を用いた配信システムの構成の一例を示すブロック図である。

【符号の説明】

- 10, 200 … クラスタサーバ装置
- 20 … 負荷分散サーバ
- 21 … クラスタ制御部
- 30, 300 … コンテンツサーバ
- 31, 301 … ストリームデータ蓄積部
- 32, 302 … 読み出し部
- 33, 221 ~ 22n, 303 … キャッシュ
- 34, 231 ~ 23n, 304 … 配信部
- 41 ~ 4n, 401 ~ 40n … 端末装置
- 51 ~ 53, 501, 503 … ネットワーク
- 101 ~ 10n, 201 ~ 20n … キャッシュサーバ

111 ~ 11n … セッション管理部

121 ~ 12n … データ配信部

131 ~ 13n … 配信不能検知部

141 ~ 14n … ストリームデータ管理部

151 ~ 15n … ストリームデータ記憶部

161 ~ 16n … 要求転送部

171 ~ 17n … キャッシュ制御部

181 ~ 18n … 配信復旧部

211 ~ 21n … 受信部

10 531 ~ 53n … セッション管理テーブル

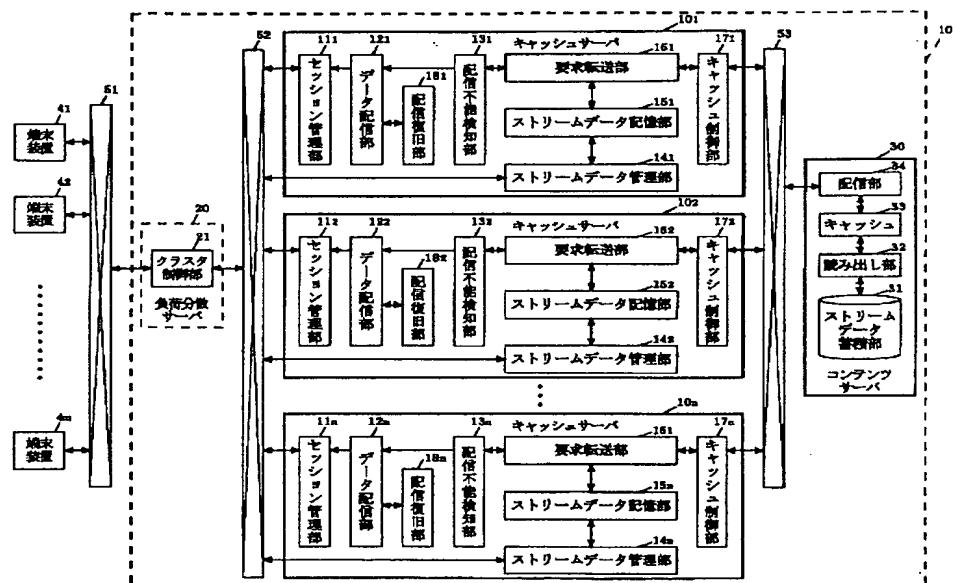
541 ~ 54n … テーブル境界値

601 ~ 60p, 701 ~ 704 … ストリームデータ

611 ~ 61q … パケット

621 ~ 62q … パケット識別子

【図 1】



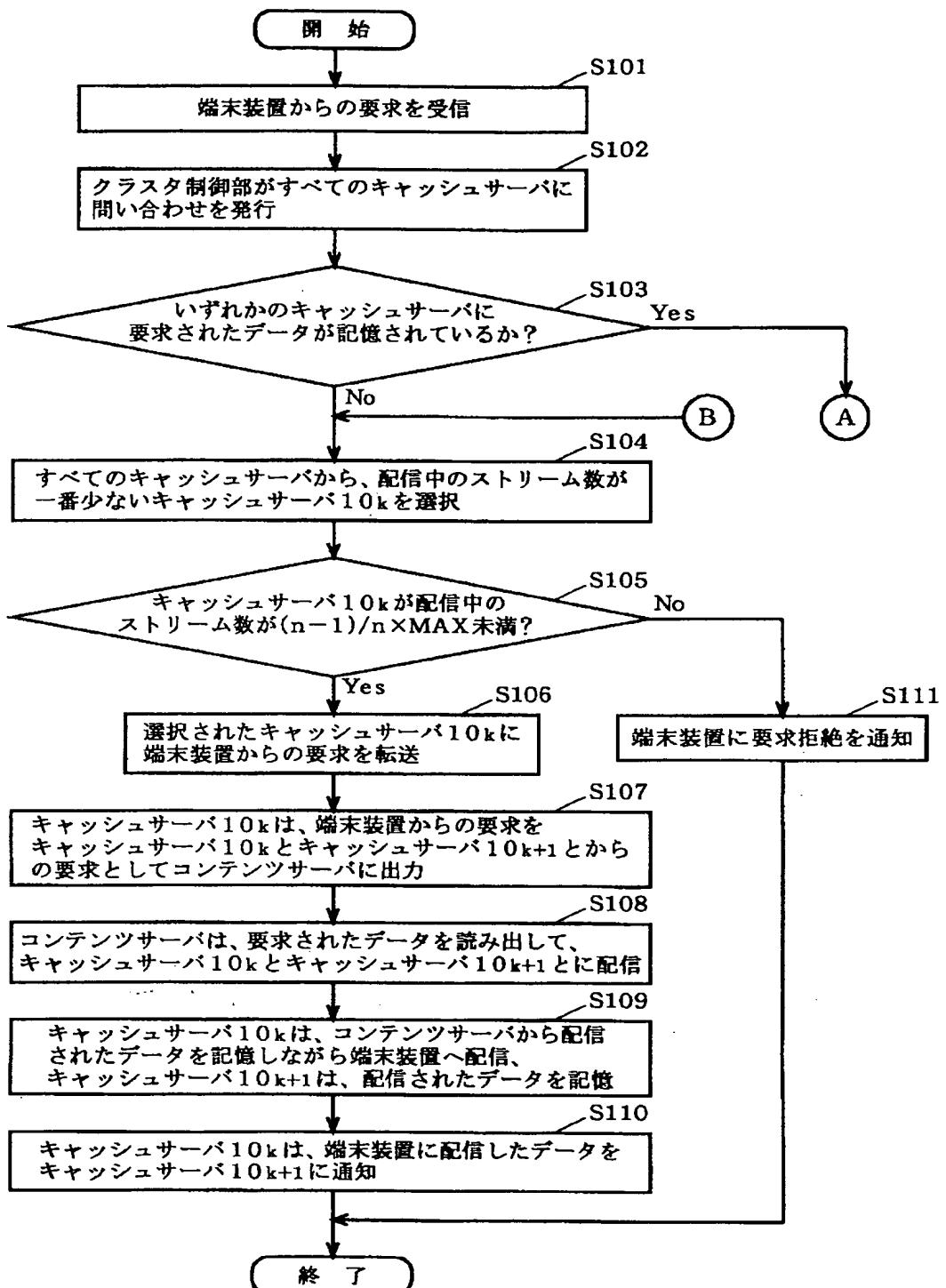
【図 4】

ストリームデータ	ストリームデータ	ストリームデータ
data1	11xx1~11yy2	xx:xx:xx
data32	123xx2~11yy3	xx:xx:xx
data13	0001x~000yz	xx:xx:xx
data24	xxxxx~122xxx	xx:xx:xx
data104	8900xx~8901xx	xx:xx:xx
.....	.....	.....

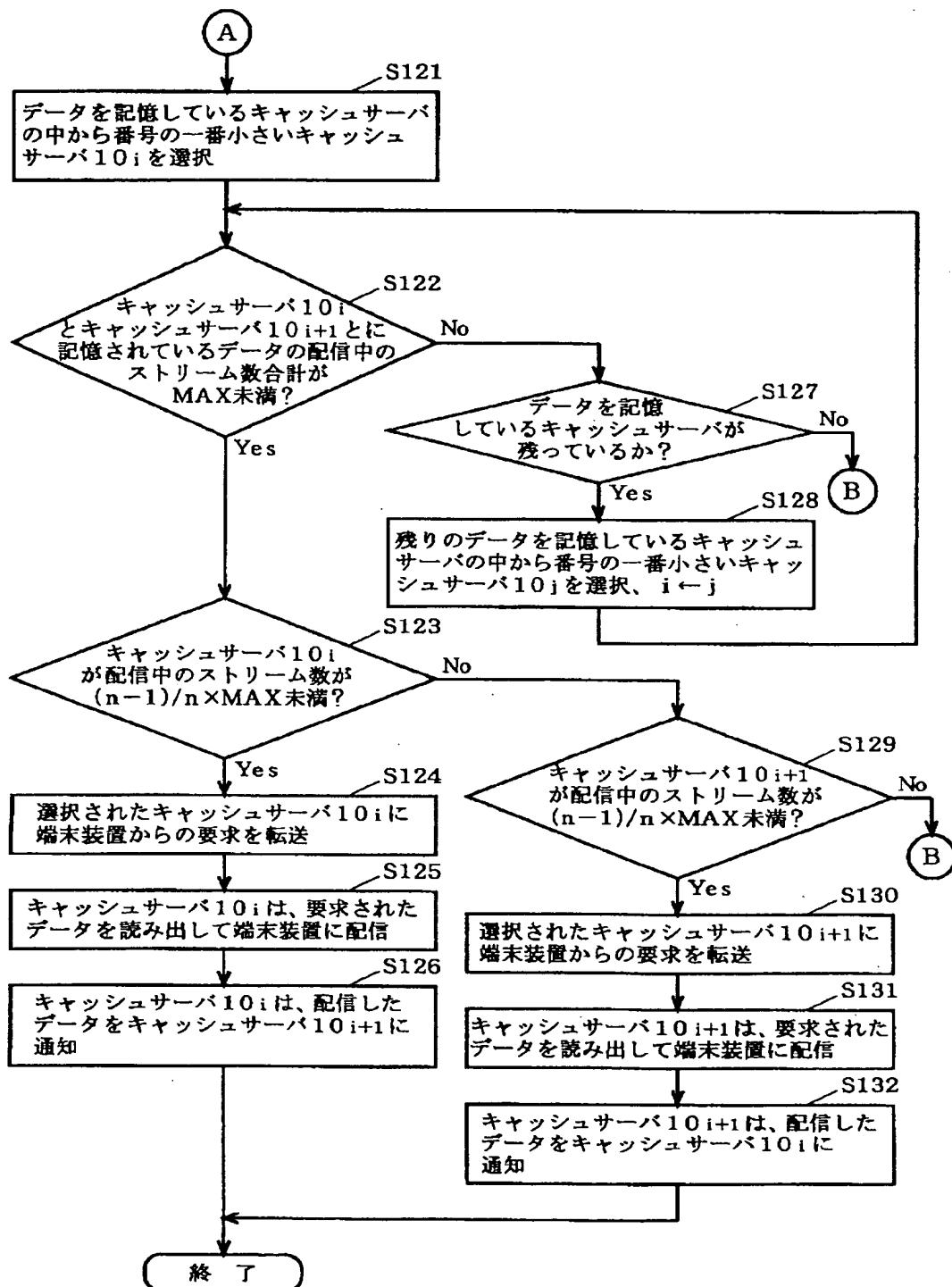
【図 5】

ストリームデータ	ストリームデータ	ストリームデータ	ストリームデータ
192.168.1.11	10002	xxxxx	2
192.168.1.12	10013	yyyyy	11
192.168.1.52	10015	zzz	21
192.168.1.31	10220	yyyyy	19
.....	.....	.....	.....
192.168.1.21	11011	xxxxx	28

[図2]



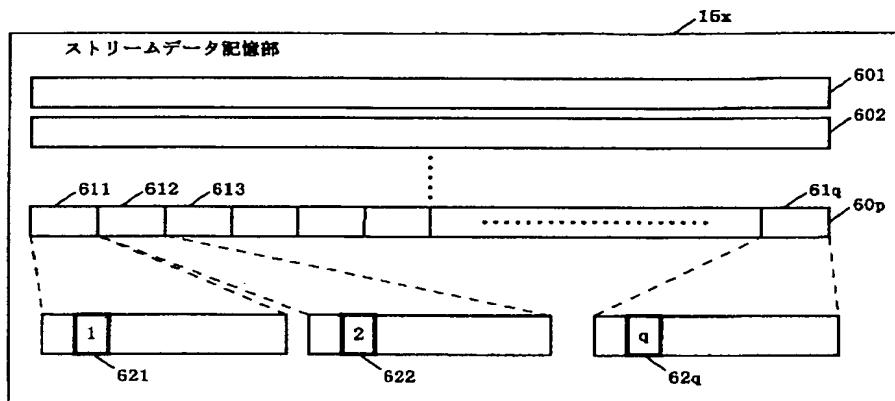
【図3】



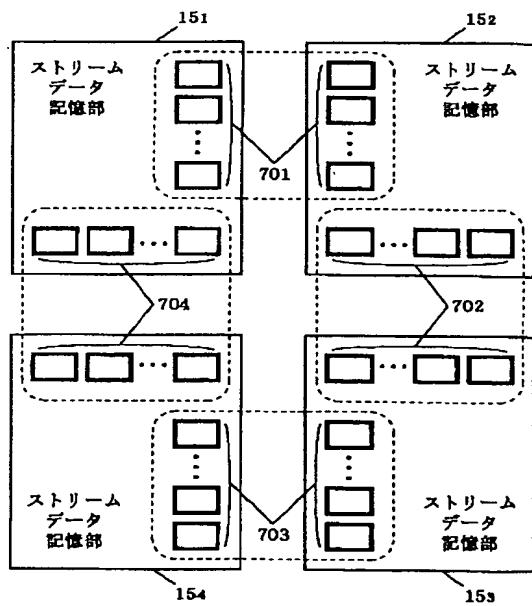
【図 6】

ストリームデータ記憶部	セッション管理部	ストリームデータ記憶部	セッション管理部
1	x	100	500
2	o	120	1000
3	o	10	500
4	x	150	200
.....	.....	.....	.....
n	x	200	500

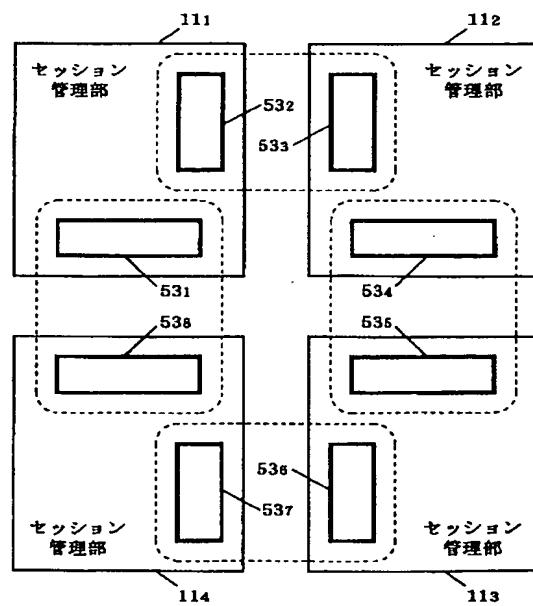
【図 7】



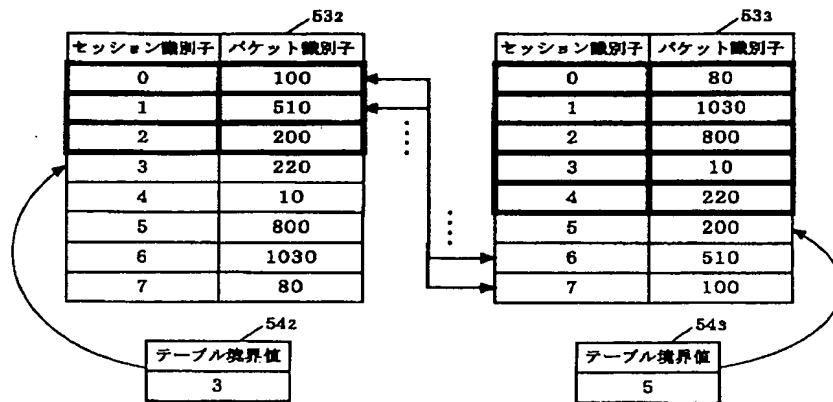
【図 8】



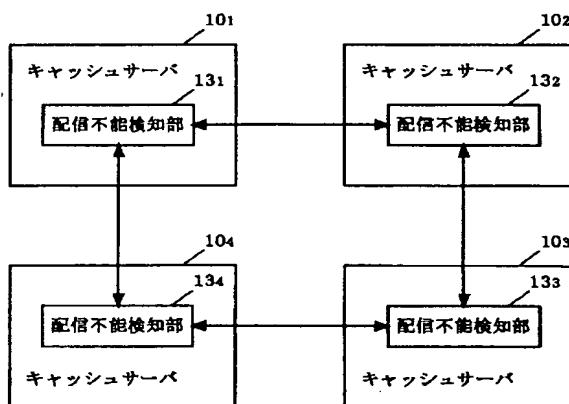
【図 9】



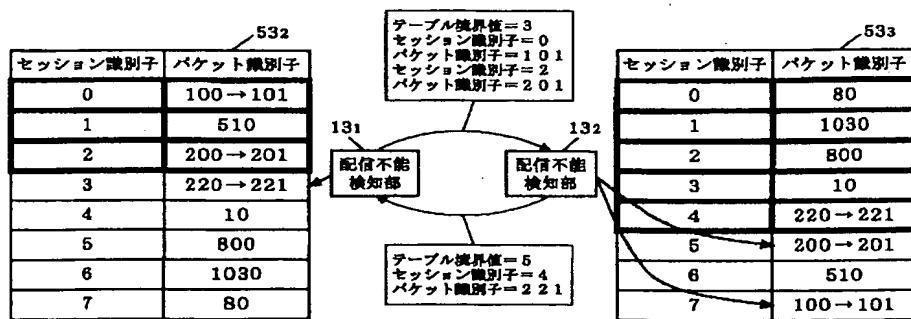
【図 10】



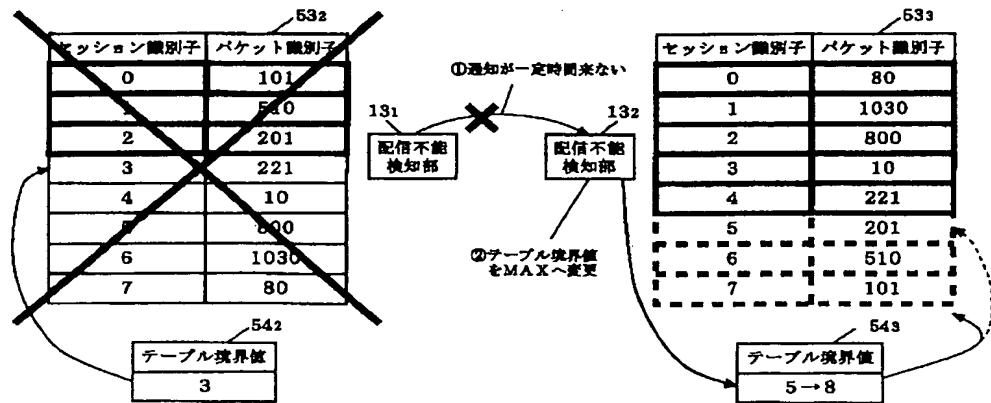
【図 11】



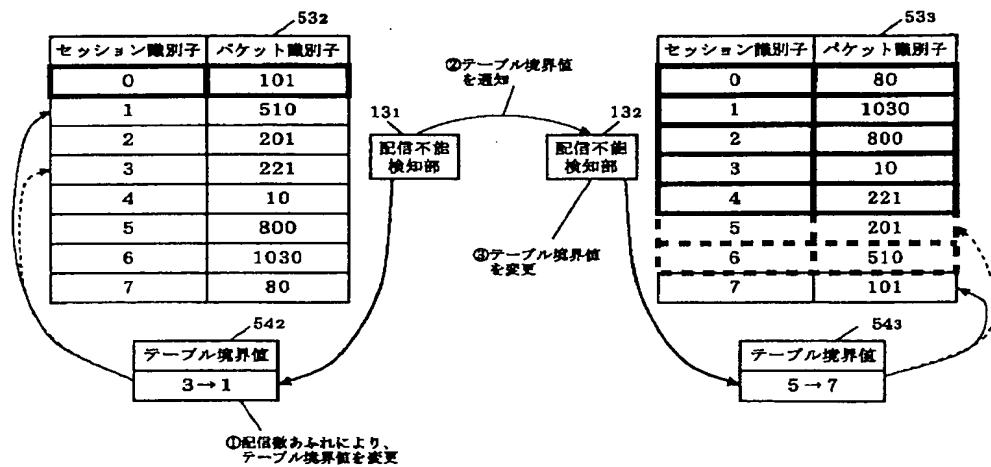
【図 12】



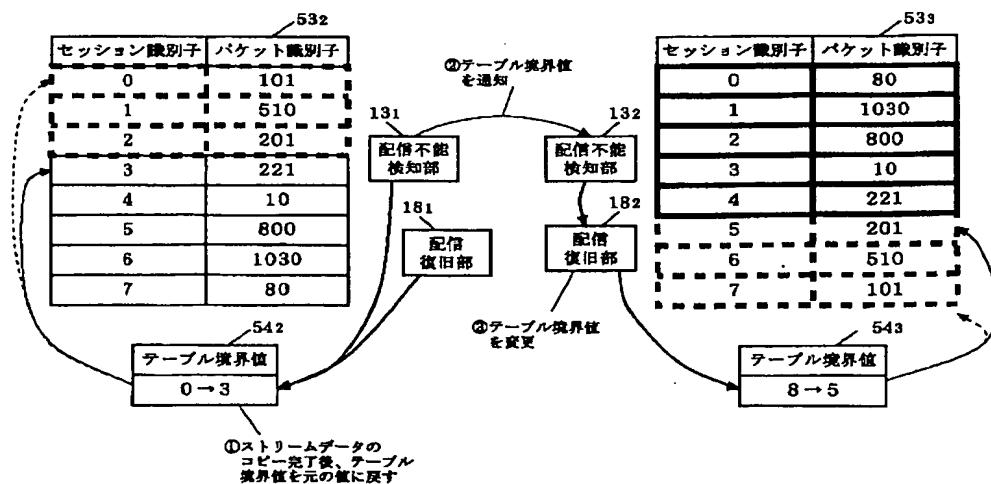
【図13】



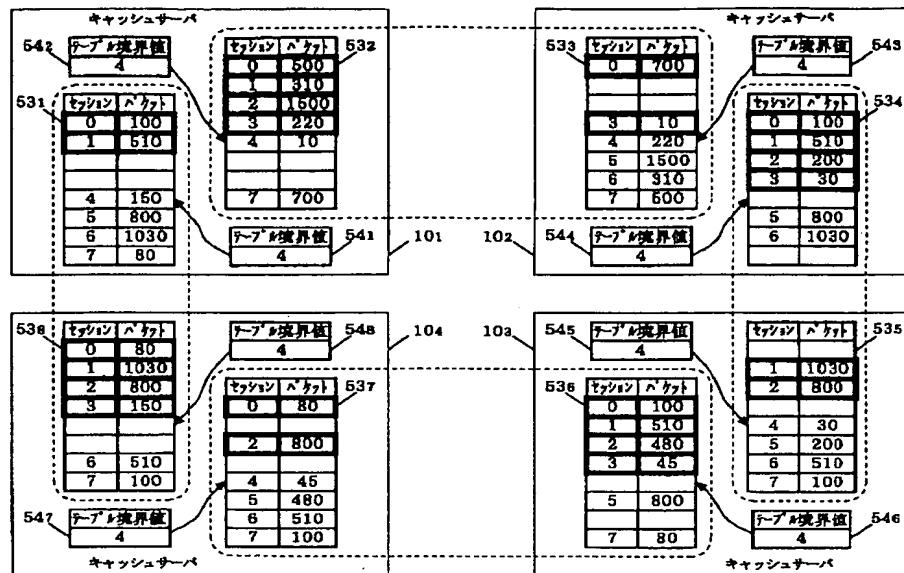
【図14】



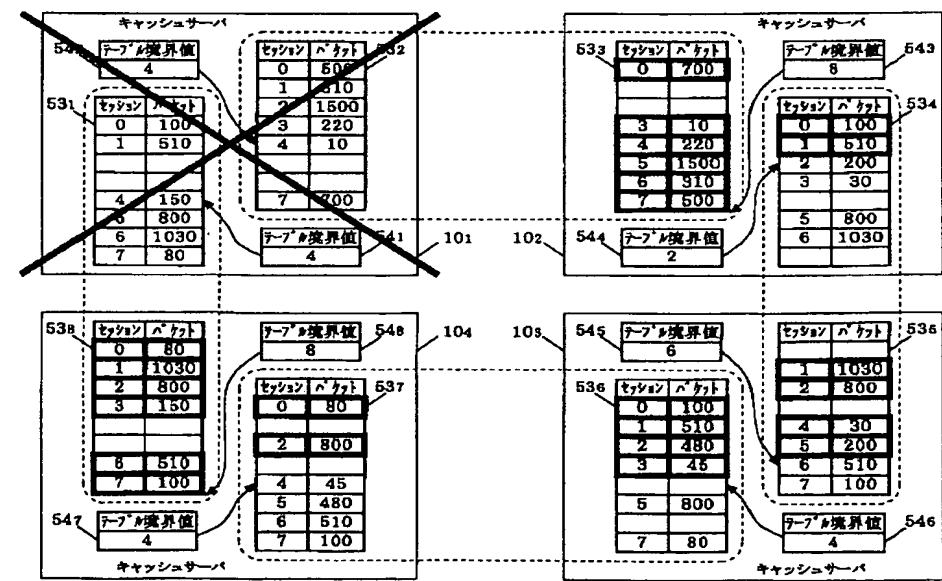
【図15】



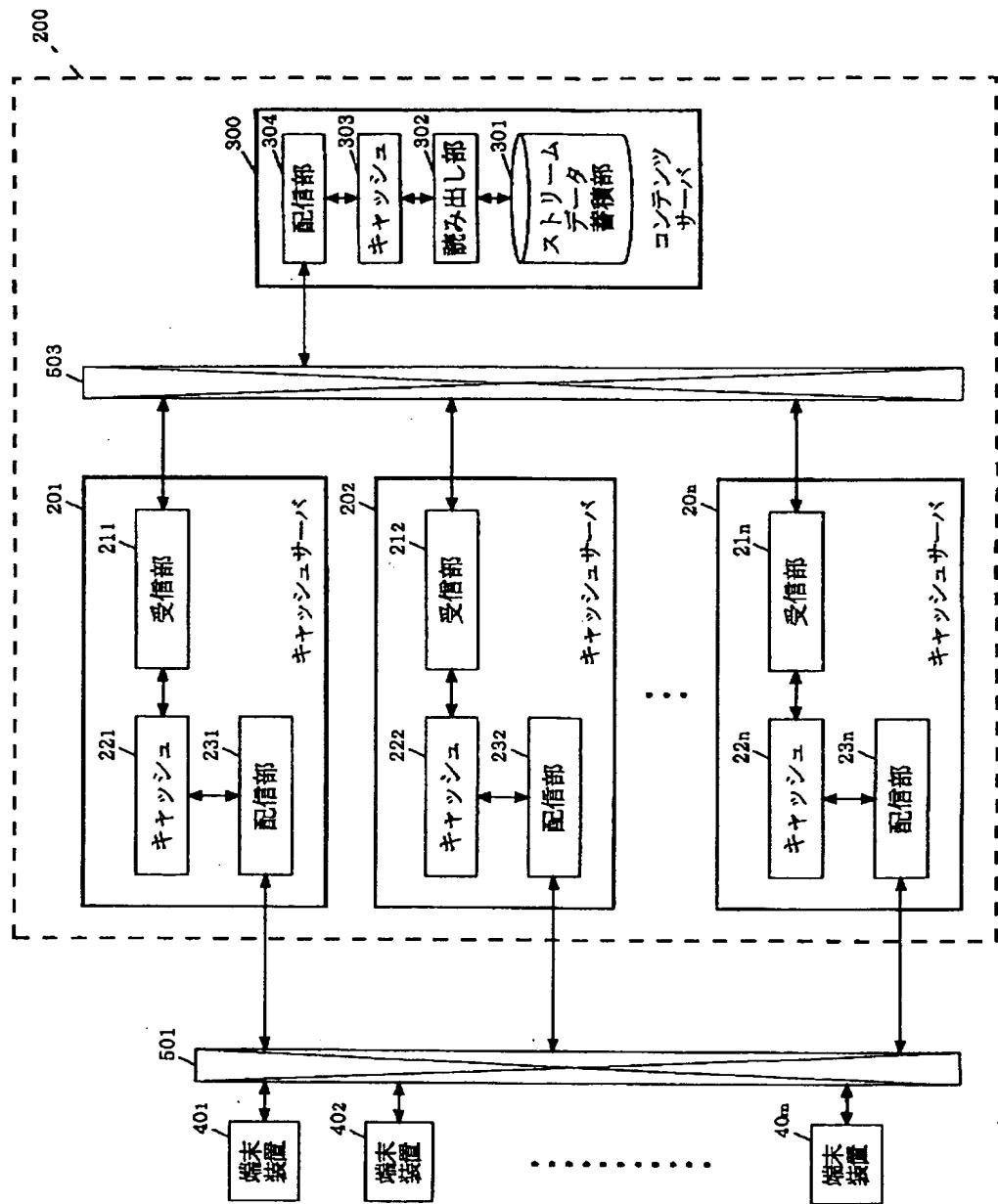
【図 16】



【図 17】



【図18】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**